

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2001, 14:37:17 ; Search time 41.92 Seconds  
(without alignments) 1965.609 Million cell updates/sec

Title: US-09-502-698-1

Perfect score: 6990

Sequence: 1 MWCRRYPTPTQSELDSD.....EEAVAIKRTKGVTLWVLS 1373

Scoring table:

BLOSUM62

Gap 10.0 , Gapext 0.5

Searched: 412676 seqs, 60623988 residues

Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A-Geneseq\_0601:\*

- 1: /SIDS8/gcgdata/geneseq/geneseqp/AA1980.DAT:\*
- 2: /SIDS8/gcgdata/geneseq/geneseqp/AA1981.DAT:\*
- 3: /SIDS8/gcgdata/geneseq/geneseqp/AA1982.DAT:\*
- 4: /SIDS8/gcgdata/geneseq/geneseqp/AA1983.DAT:\*
- 5: /SIDS8/gcgdata/geneseq/geneseqp/AA1984.DAT:\*
- 6: /SIDS8/gcgdata/geneseq/geneseqp/AA1985.DAT:\*
- 7: /SIDS8/gcgdata/geneseq/geneseqp/AA1986.DAT:\*
- 8: /SIDS8/gcgdata/geneseq/geneseqp/AA1987.DAT:\*
- 9: /SIDS8/gcgdata/geneseq/geneseqp/AA1988.DAT:\*
- 10: /SIDS8/gcgdata/geneseq/geneseqp/AA1989.DAT:\*
- 11: /SIDS8/gcgdata/geneseq/geneseqp/AA1990.DAT:\*
- 12: /SIDS8/gcgdata/geneseq/geneseqp/AA1991.DAT:\*
- 13: /SIDS8/gcgdata/geneseq/geneseqp/AA1992.DAT:\*
- 14: /SIDS8/gcgdata/geneseq/geneseqp/AA1993.DAT:\*
- 15: /SIDS8/gcgdata/geneseq/geneseqp/AA1994.DAT:\*
- 16: /SIDS8/gcgdata/geneseq/geneseqp/AA1995.DAT:\*
- 17: /SIDS8/gcgdata/geneseq/geneseqp/AA1996.DAT:\*
- 18: /SIDS8/gcgdata/geneseq/geneseqp/AA1997.DAT:\*
- 19: /SIDS8/gcgdata/geneseq/geneseqp/AA1998.DAT:\*
- 20: /SIDS8/gcgdata/geneseq/geneseqp/AA1999.DAT:\*
- 21: /SIDS8/gcgdata/geneseq/geneseqp/AA2000.DAT:\*
- 22: /SIDS8/gcgdata/geneseq/geneseqp/AA2001.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score  | Query Match | Length | DB ID | Description |
|------------|--------|-------------|--------|-------|-------------|
| 1          | 6990   | 100.0       | 1373   | 20    | AAV04730    |
| 2          | 6977   | 99.8        | 2000   | 20    | AAV04732    |
| 3          | 6921   | 99.0        | 2070   | 20    | AAV04733    |
| 4          | 6914.5 | 98.9        | 2037   | 21    | AAV53753    |
| 5          | 5085   | 72.7        | 1005   | 20    | AAV04731    |
| 6          | 4080.5 | 58.4        | 856    | 21    | AAV01383    |
| 7          | 3840   | 54.9        | 763    | 20    | AAV04741    |
| 8          | 2981   | 42.6        | 1239   | 20    | AAV04734    |
| 9          | 2457.5 | 35.2        | 1881   | 20    | AAV24025    |
| 10         | 1501   | 21.5        | 318    | 20    | AAV74150    |
| 11         | 1266   | 18.1        | 251    | 20    | AAV04738    |

|    |       |      |      |    |          |                    |
|----|-------|------|------|----|----------|--------------------|
| 12 | 1228  | 17.6 | 272  | 20 | AAV04739 | PDZ domain-contain |
| 13 | 1228  | 17.6 | 319  | 20 | AAV04740 | PDZ domain-contain |
| 14 | 928   | 13.3 | 206  | 20 | AAV74151 | Human prostate tum |
| 15 | 814   | 11.6 | 590  | 20 | AAV04736 | PDZ domain-contain |
| 16 | 583   | 8.3  | 632  | 21 | AAV6689  | Membrane-bound pro |
| 17 | 583   | 8.3  | 632  | 22 | AAV87545 | Human PRO1136, HO  |
| 18 | 583   | 8.3  | 632  | 22 | AAV65212 | Human PRO1136 (UNQ |
| 19 | 580   | 8.3  | 632  | 21 | AAV01385 | Neuron-associated  |
| 20 | 509   | 7.3  | 2485 | 21 | AAV19343 | Amino acid sequenc |
| 21 | 508   | 7.3  | 674  | 20 | AAV92953 | Fly transductosome |
| 22 | 500.5 | 7.2  | 2466 | 16 | AAV71498 | Human protein tyro |
| 23 | 500.5 | 7.2  | 2466 | 19 | AAV5999  | Intracellular prot |
| 24 | 500.5 | 7.2  | 2466 | 21 | AAV90272 | Human PRP11 phosph |
| 25 | 418.5 | 6.0  | 1161 | 21 | AAV03832 | Activin receptor b |
| 26 | 409   | 5.9  | 1112 | 21 | AAV03833 | Activin receptor b |
| 27 | 399   | 5.7  | 817  | 19 | AAV48101 | Human discs large  |
| 28 | 399   | 5.7  | 817  | 20 | AAV30137 | Amino acid sequenc |
| 29 | 399   | 5.7  | 849  | 19 | AAV48102 | Human discs large  |
| 30 | 398.5 | 5.7  | 1277 | 21 | AAV03149 | Rat synaptic scaff |
| 31 | 385   | 5.5  | 767  | 21 | AAV22138 | Human post-synapti |
| 32 | 376   | 5.4  | 344  | 22 | AAV55834 | PDZ encoded domain |
| 33 | 376   | 5.4  | 344  | 22 | AAV57623 | Human post-synapti |
| 34 | 376   | 5.4  | 344  | 22 | AAV56035 | Human post-synapti |
| 35 | 372   | 5.3  | 882  | 18 | AAV34662 | Partial PSD-93 pro |
| 36 | 359   | 5.1  | 1113 | 22 | AAV64404 | Amino acid sequenc |
| 37 | 357   | 5.1  | 1112 | 19 | AAV73061 | Rat GRIP. Rattus   |
| 38 | 347   | 5.0  | 744  | 21 | AAV71410 | PDZ domain-compris |
| 39 | 339.5 | 4.9  | 396  | 18 | AAV34666 | Partial PSD-95 pro |
| 40 | 339.5 | 4.9  | 1049 | 19 | AAV73062 | Rat GRIP2. Rattus  |
| 41 | 334.5 | 4.8  | 610  | 17 | AAV94649 | Human Fas-associat |
| 42 | 315.5 | 4.5  | 414  | 22 | AAV36687 | Mammalian two-hydr |
| 43 | 303.5 | 4.3  | 227  | 21 | AAV53276 | Human colon cancer |
| 44 | 303.5 | 4.3  | 419  | 22 | AAV64997 | Human secreted pro |
| 45 | 298   | 4.3  | 207  | 22 | AAV57625 | Human homolog of D |

#### ALIGNMENTS

RESULT 1

ID AAV04730 standard; Protein; 1373 AA.

XX AC AAV04730;

XX DT 06-JUL-1999 (first entry)

XX DE Protein containing PDZ domain from clone 38-2-1.

XX DE PDZ domain; gene expression; human umbilical vascular endothelial cell;

KW HUVEC; stimulation; tumour necrosis factor; TNF; protein binding;

KW cell; proliferation disorder; cancer.

XX OS Homo sapiens.

XX PN W09907846-A1.

PD 18-FEB-1999.

XX PF 12-AUG-1998; 98WO-JP03603.

XX PR 19-JUN-1998; 98UP-0189944.

XX PR 12-AUG-1997; 97JP-0230356.

XX PA (CHUG-) CHUGAI RES INST MOLECULAR MEDICINE INC.

XX PI Funahashi S, Miyata S;

XX DR WPI: 1999-167423/14;

XX DR N-PSDB: AAV29908.

XX PT Protein containing PDZ domain, whose expression is enhanced by TNF

PT stimulation - plays an important role in protein/protein

PT Interactions and is used for screening for proteins for use in  
 treatment of cell proliferation disorders such as cancer

XX Claim 1, Page 77-84; 240pp; Japanese.

CC This sequence represents a new protein containing a PDZ domain encoded  
 CC by the clone 38-2-1, whose expression in human umbilical vascular  
 CC endothelial cells (HUVEC) is enhanced by stimulation with tumour necrosis  
 CC factor (TNF) alpha. The new protein is used to identify proteins which  
 CC bind to it (particularly to the PDZ domain) and the genes encoding them,  
 CC for use in the treatment of cell proliferation disorders such as cancer.

XX Sequence 1373 AA;

Query Match Best Local Similarity 100.0%; Score 6990; DB 20; Length 1373;  
 Matches 1373; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MYCCRRVPTTQSELDLDCIDIELTERPHVDLGEFTGSEDEDPVLAAMDAGOSTEEV 60  
 Db 1 MYCCRRVPTTQSELDLDCIDIELTERPHVDLGEFTGSEDEDPVLAAMDAGOSTEEV 60  
 QY 61 QAPLAMEAGQHTLELESGKGLGFSIIDYODPDPASTVITISLVPGGIAEKDGLLP 120  
 Db 61 QAPLAMEAGQHTLELESGKGLGFSIIDYODPDPASTVITISLVPGGIAEKDGLLP 120  
 QY 121 GDRIMEFVNDVLENSLEAEALGAPSGTVRIGVAKPLPLSPGEGYVSKEDSFLXP 180  
 Db 121 GDRIMEFVNDVLENSLEAEALGAPSGTVRIGVAKPLPLSPGEGYVSKEDSFLXP 180  
 QY 121 GDRIMEFVNDVLENSLEAEALGAPSGTVRIGVAKPLPLSPGEGYVSKEDSFLXP 180  
 Db 121 GDRIMEFVNDVLENSLEAEALGAPSGTVRIGVAKPLPLSPGEGYVSKEDSFLXP 180  
 QY 181 HSCBAGLADKPLFRADIALVGTNDADLVDESTRESPYSPENDSIYSTQASISLHGSSC 240  
 Db 181 HSCBAGLADKPLFRADIALVGTNDADLVDESTRESPYSPENDSIYSTQASISLHGSSC 240  
 QY 181 HSCBAGLADKPLFRADIALVGTNDADLVDESTRESPYSPENDSIYSTQASISLHGSSC 240  
 Db 181 HSCBAGLADKPLFRADIALVGTNDADLVDESTRESPYSPENDSIYSTQASISLHGSSC 240  
 QY 241 GDGLYGSLSPPKVDIENSCDPLDLHMSLELYTONLEODEPDPVDISMGRAS 300  
 Db 241 GDGLYGSLSPPKVDIENSCDPLDLHMSLELYTONLEODEPDPVDISMGRAS 300  
 QY 301 GFTINDYTPANAIDQOXCENTIIVWTESHPSEVISSAELPSVLPDSAGKSEHLEQSS 360  
 Db 301 GFTINDYTPANAIDQOXCENTIIVWTESHPSEVISSAELPSVLPDSAGKSEHLEQSS 360  
 QY 301 GFTINDYTPANAIDQOXCENTIIVWTESHPSEVISSAELPSVLPDSAGKSEHLEQSS 360  
 Db 301 GFTINDYTPANAIDQOXCENTIIVWTESHPSEVISSAELPSVLPDSAGKSEHLEQSS 360  
 QY 361 LACNAECYMLONVSKESERTINTAKNSLGKTVSANKDGLMVSITIHGAISRDR 420  
 Db 361 LACNAECYMLONVSKESERTINTAKNSLGKTVSANKDGLMVSITIHGAISRDR 420  
 QY 361 LACNAECYMLONVSKESERTINTAKNSLGKTVSANKDGLMVSITIHGAISRDR 420  
 Db 361 LACNAECYMLONVSKESERTINTAKNSLGKTVSANKDGLMVSITIHGAISRDR 420  
 QY 421 IATGDCIISTEESTISTVNTQARAMLRRHSILIGPDITVTPAHEEFTISIGQOSGR 480  
 Db 421 IATGDCIISTEESTISTVNTQARAMLRRHSILIGPDITVTPAHEEFTISIGQOSGR 480  
 QY 421 IATGDCIISTEESTISTVNTQARAMLRRHSILIGPDITVTPAHEEFTISIGQOSGR 480  
 Db 421 IATGDCIISTEESTISTVNTQARAMLRRHSILIGPDITVTPAHEEFTISIGQOSGR 480  
 QY 481 VMAIDFSSYTGRIPELPERESEGSELTONTAYSNMNPRLVELMRPESKISGISIV 540  
 Db 481 VMAIDFSSYTGRIPELPERESEGSELTONTAYSNMNPRLVELMRPESKISGISIV 540  
 QY 481 VMAIDFSSYTGRIPELPERESEGSELTONTAYSNMNPRLVELMRPESKISGISIV 540  
 Db 481 VMAIDFSSYTGRIPELPERESEGSELTONTAYSNMNPRLVELMRPESKISGISIV 540  
 QY 541 GGRMGSRSLSGEYWRGIFIKHVEDSPAGKNGTLKPGDRIVEAPSOSESEPEKAPLCV 600  
 Db 541 GGRMGSRSLSGEYWRGIFIKHVEDSPAGKNGTLKPGDRIVEAPSOSESEPEKAPLCV 600  
 QY 541 GGRMGSRSLSGEYWRGIFIKHVEDSPAGKNGTLKPGDRIVEAPSOSESEPEKAPLCV 600  
 Db 541 GGRMGSRSLSGEYWRGIFIKHVEDSPAGKNGTLKPGDRIVEAPSOSESEPEKAPLCV 600  
 QY 601 PPPPSAFEMGSDHTOSSAKISODVDKEDDEFGYSKNTIRERCTLTGELHMLEKGH 660  
 Db 601 PPPPSAFEMGSDHTOSSAKISODVDKEDDEFGYSKNTIRERCTLTGELHMLEKGH 660  
 QY 601 PPPPSAFEMGSDHTOSSAKISODVDKEDDEFGYSKNTIRERCTLTGELHMLEKGH 660  
 Db 601 PPPPSAFEMGSDHTOSSAKISODVDKEDDEFGYSKNTIRERCTLTGELHMLEKGH 660  
 QY 661 SGLGLSLAGNKRDRMSYFIVGIDPNGAAGKGRLOJIADELLENGOITLYGRSHQNSSI 720  
 Db 661 SGLGLSLAGNKRDRMSYFIVGIDPNGAAGKGRLOJIADELLENGOITLYGRSHQNSSI 720  
 QY 661 SGLGLSLAGNKRDRMSYFIVGIDPNGAAGKGRLOJIADELLENGOITLYGRSHQNSSI 720  
 Db 661 SGLGLSLAGNKRDRMSYFIVGIDPNGAAGKGRLOJIADELLENGOITLYGRSHQNSSI 720  
 QY 721 IKCAPSRVYKIIIFRNKDAVQMAVCPGNAVEPLPSNSENIKNKTEPTVTTSDAAVLS 780  
 Db 721 IKCAPSRVYKIIIFRNKDAVQMAVCPGNAVEPLPSNSENIKNKTEPTVTTSDAAVLS 780  
 QY 721 IKCAPSRVYKIIIFRNKDAVQMAVCPGNAVEPLPSNSENIKNKTEPTVTTSDAAVLS 780  
 Db 721 IKCAPSRVYKIIIFRNKDAVQMAVCPGNAVEPLPSNSENIKNKTEPTVTTSDAAVLS 780  
 QY 781 FKNVQHLEPRDGGGLIAISEEDTSGVITIKSTEHGVAATDGRKLAVGDOIILAVDEIV 840  
 Db 781 FKNVQHLEPRDGGGLIAISEEDTSGVITIKSTEHGVAATDGRKLAVGDOIILAVDEIV 840

Db 781 FKNVQHLEPRDGGGLIAISEEDTSGVITIKSTEHGVAATDGRKLAVGDOIILAVDEIV 840  
 QY 841 VGYPIEKRTISILKTAKTMTKTIHAENPDSQAVNSAGACGEKKNSSQSLMVFQSGSPE 900  
 Db 841 VGYPIEKRTISILKTAKTMTKTIHAENPDSQAVNSAGACGEKKNSSQSLMVFQSGSPE 900  
 QY 901 PESTINSTRSSTPIAFADPATCPRIPECETTERISKRTGLGSIYGSPTLGAFTIH 960  
 Db 901 PESTINSTRSSTPIAFADPATCPRIPECETTERISKRTGLGSIYGSPTLGAFTIH 960  
 QY 901 PESTINSTRSSTPIAFADPATCPRIPECETTERISKRTGLGSIYGSPTLGAFTIH 960  
 Db 901 PESTINSTRSSTPIAFADPATCPRIPECETTERISKRTGLGSIYGSPTLGAFTIH 960  
 QY 961 EYEBGAACKDGRUWAGDQILEVNGIDIRKATHDEAINVLRQTRQVRLTYRDEAPYKE 1020  
 Db 961 EYEBGAACKDGRUWAGDQILEVNGIDIRKATHDEAINVLRQTRQVRLTYRDEAPYKE 1020  
 QY 961 EYEBGAACKDGRUWAGDQILEVNGIDIRKATHDEAINVLRQTRQVRLTYRDEAPYKE 1020  
 Db 961 EYEBGAACKDGRUWAGDQILEVNGIDIRKATHDEAINVLRQTRQVRLTYRDEAPYKE 1020  
 QY 1021 EEWCDTLTIELOKKPGKGLSIVGKRDNTGVFSVDIYKGIADPDGRLLOGDILLVNG 1080  
 Db 1021 EEWCDTLTIELOKKPGKGLSIVGKRDNTGVFSVDIYKGIADPDGRLLOGDILLVNG 1080  
 QY 1021 EEWCDTLTIELOKKPGKGLSIVGKRDNTGVFSVDIYKGIADPDGRLLOGDILLVNG 1080  
 Db 1021 EEWCDTLTIELOKKPGKGLSIVGKRDNTGVFSVDIYKGIADPDGRLLOGDILLVNG 1080  
 QY 1081 EDVRNASQEAVALKLSIGLTVTEVGRIRKAGPFSERRPQTSQVSEGLSSTFPLSG 1140  
 Db 1081 EDVRNASQEAVALKLSIGLTVTEVGRIRKAGPFSERRPQTSQVSEGLSSTFPLSG 1140  
 QY 1081 EDVRNASQEAVALKLSIGLTVTEVGRIRKAGPFSERRPQTSQVSEGLSSTFPLSG 1140  
 Db 1081 EDVRNASQEAVALKLSIGLTVTEVGRIRKAGPFSERRPQTSQVSEGLSSTFPLSG 1140  
 QY 1141 SSTSSELESSKKNALASEIOGLRTVEKKKPTDSLGISAGVSGPLGDVPIITAMMP 1200  
 Db 1141 SSTSSELESSKKNALASEIOGLRTVEKKKPTDSLGISAGVSGPLGDVPIITAMMP 1200  
 QY 1141 SSTSSELESSKKNALASEIOGLRTVEKKKPTDSLGISAGVSGPLGDVPIITAMMP 1200  
 Db 1141 SSTSSELESSKKNALASEIOGLRTVEKKKPTDSLGISAGVSGPLGDVPIITAMMP 1200  
 QY 1201 TGVAAGTQKLRVGRIVYICGSTEGMTHTQAVNMLKKNASGSIEMQVYAGGVGVTGHH 1260  
 Db 1201 TGVAAGTQKLRVGRIVYICGSTEGMTHTQAVNMLKKNASGSIEMQVYAGGVGVTGHH 1260  
 QY 1201 TGVAAGTQKLRVGRIVYICGSTEGMTHTQAVNMLKKNASGSIEMQVYAGGVGVTGHH 1260  
 Db 1201 TGVAAGTQKLRVGRIVYICGSTEGMTHTQAVNMLKKNASGSIEMQVYAGGVGVTGHH 1260  
 QY 1261 QEPASSLSFTGLTSTIFDODLGPPOCKSTTLERGPDGLGFSIVGSGSPHGLPIYVK 1320  
 Db 1261 QEPASSLSFTGLTSTIFDODLGPPOCKSTTLERGPDGLGFSIVGSGSPHGLPIYVK 1320  
 QY 1261 QEPASSLSFTGLTSTIFDODLGPPOCKSTTLERGPDGLGFSIVGSGSPHGLPIYVK 1320  
 Db 1261 QEPASSLSFTGLTSTIFDODLGPPOCKSTTLERGPDGLGFSIVGSGSPHGLPIYVK 1320  
 QY 1321 TVPKKGAASEGRKLRGDOILIAVNGOSLEGVTHEEAVALIKRTGCTIMWLIS 1373  
 Db 1321 TVPKKGAASEGRKLRGDOILIAVNGOSLEGVTHEEAVALIKRTGCTIMWLIS 1373

RESULT 2  
 ID AAY04732  
 AAAY04732 standard; Protein: 2000 AA.  
 AC AAY04732;  
 DT 06-JUL-1999 (first entry)  
 XX  
 DE Protein containing PDZ domain from clone 38-2-1a.  
 XX  
 KW PDZ domain; gene expression; human umbilical vascular endothelial cell;  
 KW HUVEC; stimulation; tumour necrosis factor; TNF; protein binding;  
 KW cell; proliferation disorder; cancer.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WC0907846-A1.  
 PD 18-FEB-1999.  
 XX  
 PF 12-AUG-1998; 98WO-JP03603.  
 XX  
 PR 19-JUN-1998; 98JP-0189944.  
 PR 12-AUG-1997; 97JP-0230356.  
 XX  
 PA (CHUG-) CHUGAI RES INST MOLECULAR MEDICINE INC.  
 PI Funahashi S, Miyata S;  
 DR WPI: 1999-167423/14.  
 DR N-PDB; AAX29908.  
 XX  
 PT Protein containing PDZ domain, whose expression is enhanced by TNF  
 stimulation - plays an important role in protein/protein

Interactions and is used for screening for proteins for use in treatment of cell proliferation disorders such as cancer

Claim 1: Page 158-167; 240pp; Japanese.

This sequence represents a new protein containing a PDZ domain encoded by the clone 38-2-1a, whose expression in human umbilical vascular endothelial cells (HUVEC) is enhanced by stimulation with tumour necrosis factor (TNF) alpha. The new protein is used to identify proteins which bind to it (particularly to the PDZ domains) and the genes encoding them, for use in the treatment of cell proliferation disorders such as cancer.

Sequence 2000 AA:

Query Match 99.8%; Score 6977; DB 20; Length 2000;  
Best Local Similarity: 99.9%; Pred. No. 0;  
Matches 1371; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

QY 1 MWCCRRYPTPTQSELDSDLCDELTEKPHVDLGEFTGSSSTEDPVLAMTACGSTERY 60
DB 628 mwccrryptptqseldsdldclteltekhvdlgeftgsssepedpvlamtacgsteev 687
QY 61 QAPLAMEAGIOHIELEKSGKLGSIIDYODIPASTVITIRSLVPGIAEKDGRLLP 120
DB 668 qaplamwagiohieleksgksgisildyodipastvilitrslvpggiaekdgrllp 747
QY 121 GDRLMFVNDVNLSSLEBAVALKAPSGTVIGAVKAPLSPGEGYSAKEDSFLYPP 180
DB 748 gdrlmfvndvnlsslebavalkapsgtvirgavkaplspgegygakedsflypp 807
QY 181 HSCENAGLADKRLFRADLALVGTNDADLVDESTFSPSPENDSTYQTASTLSHGSSC 240
DB 808 hscenagladkrlfradlaltvgtndadvdestfsspendsytqastlslhgssc 867
QY 241 GGLINVGSSLPSPPKVDIENSCDPIVLDHMSLEELYTONLERDENPVDISMGKPS 300
DB 868 gglinvgsslpsspkkvdienscdpivldhmsleelytonlerdentpvdismgkps 927
QY 301 GTTINDYTPANAIEGOYECEENTIVTESHLPESEVTSASLPVLPDSAKGSEHLLEQSS 360
DB 928 gttindytpanaiegoyecentivteshlpesevtsaslpvlpdsakgsehllqess 987
QY 361 LACNAECVWLQNVSKESFERTINIAKGNSSLGTMVSANKDGMIVRSIIHGCAISRDR 420
DB 988 lacnaecvwlqnvskesfertiniakgnsslgtmvsankdgmivrsiihgcaisrdr 1047
QY 421 IAGDCIISINEESTISVTNAQARALRRHSLIGPDKITTYVPAEHLLEFKTISLQOQSR 480
DB 1048 iagdciiisineestisvtnaqaralrrhsligpdkittyvpaehllefktislqoqsr 1107
QY 481 VVALDIFSSYTGRODPELPEREGEGESELONTAISMNORRRLVLMKPEKSLGISTIV 540
DB 1108 vvaldifssytgrodpeleperegegeeselonataysmnorrrlvmkpeksglisliv 1167
QY 541 GGRKGMSRLSNGEVMGIRIKHVLIEDSPAGKNGTLKPGRIEAPVQSSSEPEKAPLCV 600
DB 1168 ggrkgmsrlsngewmgiirkhvliledspagkngtlkpgrievapvqsssepekaplcv 1227
QY 601 PPPPSAFAMGSDHTQSSASKISQDVKDEDEFGYSWKNIIRERYGTLGELHMLEEKH 660
DB 1228 ppppsafamgshdhtqssaskisqdvkdeedfgyswknirerygtlgehlmlleekh 1287
QY 661 SGLGLSLACNKRBSRMASVITIGIDPNGAAGKGRLOIADLELTNQIILXGSHQNASST 720
DB 1288 sglglslacnkrbsrmasvitigidpngaagkgrloiadleltngiilxgshqnasst 1347
QY 721 ITCAPSKVITIFRNKDAVNOMAVCPGNAVEPSPNSSENLOKNETPPTVTSAAVDLSS 780
DB 1348 itcapskvilitfrnkdaavnqmvcpgnavepssnsenlonketpvttsaavdlss 1407
QY 781 FKNVQHLLEPKDGGGLGIAISEEDTLGVIKSLTEHGVAATDGRUKVGDQILAVDEIV 840

```

```

DB 1408 fknvqhllpkdggglgiaiseedtlsgviksltehgvaatdgrlukvgdqilavdeiv 1467
QY 841 VGYPIEKRTISLKTAKMVKTLTHAENPDQAVPSAAGASEKKNSSQSLMWPOGSGPE 900
DB 1468 vgypiekrtisllktakmvtklthae npdqavpsaagasekknssqslmwpgsgpe 1527
QY 901 PESINTSRSSPFAFADPAPCPPIPGCEITTEISKRTGSLGIVSGSPDLGAFIIFH 960
DB 1528 pesintsrsscpafadpaccpplipgcetlteiskrtgslgsvsgspdlgafiih 1587
QY 961 EYEEGAACKDGRWLWAGDOILEVNGIDLRKATHDEAINVLRQTPQVRVLTLYRDEAPYKE 1020
DB 1588 eyeeegaackdgrlwa gqdllevngidl rka thdeainvlrqtpqvrvtllyrdeapyke 1647
QY 1021 EEVCDTLTIELOKRRKGLGSLIVCKRNDTVGFVSDIYKGGIADPDGRILIGDQILLYNG 1080
DB 1648 eevcdtltie lo krrkglgslivckrndtv gfv sdi yk ggi adp dgr il ig dqi lly ng 1707
QY 1081 EDVRNASEAVNALIKCSIGTVLTLEVGRIKAGPFHSERPSQTSQVSEGLSSFFPPLSG 1140
DB 1708 edvrnas eavna l i kcs i gtv l tle v g r i k a g p f h s e r p s q t s q v s e g l s s f f p p l s g 1767
QY 1141 SSTSESLSSSKKNALASEIOLGRTVEKKKGPDSLGISLAGVSPGLGDVPIFIAMMHP 1200
DB 1768 sstseslssskknalaseiolg r tve m k k g p d s l g i s l a g v s p g l g d v p i f i a m m h p 1827
QY 1201 TGVAAQOTOKLRVGDRIVITCGISTGEMHTQAVNLKNAASGIEBQVYAGSVVYTGHH 1260
DB 1828 tgvaaqotoklrvgdri vitcgistgemhtqavnlknasgiebqv y ag s v v y t g h h 1887
QY 1261 QEPASSISFTGLTSTNIFPDOLGPPOCKSTLERKPDGLGFSIYGVGSPPGDLPITYK 1320
DB 1888 qepassisftglstsnifpdolgppockstlerkpdglgfsiygvgsppgdlpityk 1947
QY 1321 TVFAKGAASEDGRLRGDQIIIVANGOSLEGVTHEEVAAILKFTKGTVTLMVLVS 1373
DB 1948 tvfakgaase dgr l r g d q i i v a n g o s l e g v t h e e v a i l k f t k g t v t l m v l v s 2000

RESULT 3
AA104733
ID AAY04733 standard; Protein: 2070 AA.
XX
AC AAY04733;
XX
DT 06-JUL-1999 (first entry)
XX
DE Protein containing PDZ domain from clone 38-2-1b.
XX
KW PDZ domain; gene expression; human umbilical vascular endothelial cell;
KW HUVEC; stimulation; tumour necrosis factor; TNF; protein binding;
KW cell; proliferation disorder; cancer.
XX
OS Homo sapiens.
XX
PN WO9907846-A1.
XX
PD 18-FEB-1999.
XX
PF 12-AUG-1998; 98WO-TP03603.
XX
PR 19-JUN-1998; 98JP-0189944.
XX
PR 12-AUG-1997; 97JP-0230356.
XX
PA (CHUG-) CHUGAI RES INST MOLECULAR MEDICINE INC.
XX
PI Funahashi S, Miyata S;
XX
DR MPI: 1999-167423/14.
XX
DR N-PSDB; AAY29910.
XX
PT Protein containing PDZ domain, whose expression is enhanced by TNF
stimulation - plays an important role in protein/protein

```

PT Interactions and is used for screening for proteins for use in  
PT treatment of cell proliferation disorders such as cancer

PS Claim 1; Page 167-176; 240pp; Japanese.

CC This sequence represents a new protein containing a PDZ domain encoded  
CC by the clone 38-2-1b, whose expression in human umbilical vascular  
CC endothelial cells (HUVEC) is enhanced by stimulation with tumour necrosis  
CC factor (TNF) alpha. The new protein is used to identify proteins which  
CC bind to it (particularly to the PDZ domains) and the genes encoding them,  
CC for use in the treatment of cell proliferation disorders such as cancer.

SO Sequence 2070 AA:

Query Match 99.08; Score 6921; DB 20; Length 2070;  
Best Local Similarity 94.98; Pred. No. 0;  
Matches 1369; Conservative 1; Mismatches 3; Indels 70; Gaps 1;

```

QY 1 MWCRRTPPTTOSLSDLDLCLDIETKPAVDGEFISGSETPVLAAMDAGQSTEEV 60
DB mwcrrtppttgseldsldclclietkphvdgefsgsepedpvlamtdagqsteev 687
QY 61 QAPLAMEAGIOHTELEKSGGLGFSILDYQDPTDPASTVITISLVPGGAENKGRLLP 120
DB qaplameagiqhileksgkglgfsildyqdpdpastvilitrlvpggaekdgrllp 747
QY 688 qaplameagiqhileksgkglgfsildyqdpdpastvilitrlvpggaekdgrllp 180
DB 688 qaplameagiqhileksgkglgfsildyqdpdpastvilitrlvpggaekdgrllp 747
QY 121 GDRIMFNDVLENSLEAEVAKGAPSGTVRIGVAKPLPLSPGEGYSAKEDSFYYP 180
DB 121 gdrimfndvleensleaevalkgapsgtvrigvakplplspgegyssakedsflypp 807
QY 748 gdrimfndvleensleaevalkgapsgtvrigvakplplspgegyssakedsflypp 807
QY 181 HSCBAGLADKPLFRADLAVGTNDADLVDSFTESPYSPENDSIYSTQASISLHGSSC 240
DB hscbagladkplfradlavgtndadvdsftespyspendsiytqasislhggssc 867
QY 808 hscbagladkplfradlavgtndadvdsftespyspendsiytqasislhggssc 867
QY 241 GDGIVNVSSTSPSPKPVIVENSOPVLDMHSLBELTQNLERQDENTSVDSMKRPAS 300
DB 241 gdgivnvsstspspkpvivensopvldmhslbeltonlerqdentssvdsmpas 927
QY 868 gdgivnvsstspspkpvivensopvldmhslbeltonlerqdentssvdsmpas 927
QY 301 GFTINDTPNNAIHOYECEENTIVWTESHLPSEYISSAELPSVLPDAGKSEHLLPOSS 360
DB 301 gftindtpnnaihoyecentiwteshlpseyissaelpsvlpdagksehllooss 987
QY 928 gftindtpnnaihoyecentiwteshlpseyissaelpsvlpdagksehllooss 987
QY 361 LACNAECVMLQNVSKESFERTINIAKGNSSLGMTVSANKDGLMIVRSIIHGASISDRGR 420
DB 361 lacnaecvmlqnvskesfertinlakgnsslgmtvsankdglmivrsilhgasistrdgr 1047
QY 988 lacnaecvmlqnvskesfertinlakgnsslgmtvsankdglmivrsilhgasistrdgr 1047
QY 421 IATGDCILSTINEESTISVTNQAARMLRRHSLIGPDIKITVPAHEHEEKISLGGQSGR 480
DB 421 iatgdcilstineestisvtnqaarmlrrhsligpdikityvpaheheekislgqsgsr 1107
QY 1048 iatgdcilstineestisvtnqaarmlrrhsligpdikityvpaheheekislgqsgsr 1107
QY 481 VMAIDFSSYTGRIPELPEREEGEGESELQNTAYSNMNPORREVLMPREPSKLSGISIV 540
DB 481 vmaidfssytgripelpereegegeselqntaysnmwprrevelmpresklsgisiv 1167
QY 1108 vmaidfssytgripelpereegegeselqntaysnmwprrevelmpresklsgisiv 1167
QY 541 GGRGMSGRSLNNGEYMGKIFIKHVLKEDSPACKNGTLKPDRIVE-----583
DB 541 ggrgmsgrslngeymgkifikhvledspackngtlkpdrive-----583
QY 1168 ggrgmsgrslngeymgkifikhvledspackngtlkpdrive-----583
DB 1168 ggrgmsgrslngeymgkifikhvledspackngtlkpdrive-----583
QY 584 -----APQSSES 590
DB 584 -----APQSSES 590
QY 1228 lrragrpvfmvgllmrprkspjllhnllypkynfstnfpadsllqadkapsqses 1287
DB 1228 lrragrpvfmvgllmrprkspjllhnllypkynfstnfpadsllqadkapsqses 1287
QY 591 EPEKAPLCSVPPPPSAFAEMGSDHNTQSSASKISQVDKEDDEGYSWKNTIRERYGLTGE 650
DB 591 epekaplcsvppppsaafaemgshntqssaskisqvdkeddegyswkntireryglatge 1347
QY 1288 epekaplcsvppppsaafaemgshntqssaskisqvdkeddegyswkntireryglatge 1347
QY 651 LHMTELEKSGGLSLAGAKKDSRMSVFTVGIDPNGAAGKGRGLQIADLELLEINGOILY 710
DB 651 lhmteleksgglslagakkdsrmsvftvgidpngaagkgrglqiadelletleingilly 1407
QY 1348 lhmteleksgglslagakkdsrmsvftvgidpngaagkgrglqiadelletleingilly 1407
QY 711 GRSHQNASSTIKCAPSVKTIIFRNKDAVNQMAVCPGNANPELPNSNENQNKETEPTV 770
DB 711 grshqnassstikcapsvkttiifrnkdavnqmaavcpgnanpelepnsnenqnketeptv 770

```

```

DB 1408 grshqnassstikcapsvkttiifrnkdavnqmaavcpgnanpelepnsnenqnketeptv 1467
QY 771 TSDAVALDSSEFNKYOHLELPEKDOGLGIAISEEDTLGVTIKRSTHEGVAATGRLKVC 830
DB 1468 tsdaavaldssefnkyohlelpekdoglglaisseedtlgvtikrsthegvaatgtrlkvcd 1527
QY 831 QILAVDEIVVGYPIEKFISLLKTAKTAVKLTTHAENPDSQAVPSAAGAAGEKKNSSQS 890
DB 1528 qilavdeivvgypiekfisllktaktavkltthaenpdsqavpsaagaagekknssqs 1587
QY 891 LMPVQSSPPEPESIRNRSRSTPAIFASDPATCIIITGCTTIEISGRGLGISTYGS 950
DB 1588 lmpvqssppepessirnrstspafasdpatciiitgcttieleisgrglgistygs 1647
QY 951 DTLGAFIHEVEEGEAACDGRMLAGDQILEVNGIDLRKATHDEANLVROTPOVRILT 1010
DB 1648 dtlgafileveegeaacdgrmlagdqillevngidlrkatndeaanlvrotppqvrlt 1707
QY 1011 LYRDEAPYKEEVCDTLTIELQKKPKGKGLSLSTYGRKNDGVFVSDIYKGLADPDGRLL 1070
DB 1708 lydeapykeevcdtltelqkkpkkgkglslstygrkndgvfvsdiykgladpdgrll 1767
QY 1071 OGDQIILVNGEDVPMASOEAVALKCSLGMTVEVGRIRKAGPHSRRSOTSQVSEGS 1130
DB 1768 ogdqiiilvngedvpmasoeavalkcslgmtvevgrirakgphsrrsotsqvsegs 1827
QY 1131 LSSEFTPLSGSTSESLSSSKKNALASEIQGLRTVEMKKGPDSDLSIAGVGSPLGD 1190
DB 1828 lsstfplsgstseesskknalaseliglrvtemkkpdcslsagvgapldg 1887
QY 1191 VPFIAMHPTGVAAGTQKLRVGDRIYVTCGSTEGMTHTQAVNLKNASGSIEMOVVAG 1250
DB 1888 vpfiamhptgvaagtqklrvgdriyvtcgstegmthtqavnlknasgsiemovvag 1947
QY 1251 GDVSVYTGHHOEPAASSLSFTGLTSTSIPODDIGRPCKSITLERGPDGSGFSYGS 1310
DB 1948 gdvsyvtyghoePASSLSFTGLTSTSIPODDIGRPCKSITLERGPDGSGFSYGS 2007
QY 1311 PHGDLPYKTVFAKGAASEDGRKLRGDQITIAVNGQSLSEGVTHEEVAAILKRTKGTTL 1370
DB 2008 phgdlpkytvfakgaasedgrklrgdqitavngqslsegvtheevaailkrktgttlm 2067
QY 1371 VLS 1373
DB 2068 vls 2070

```

RESULT 4  
AA53753 standard; Protein; 2037 AA.  
ID AA53753;  
XX  
AC AA53753;  
XX  
DT 22-FEB-2000 (first entry)  
XX  
DE Amino acid sequence of the MMSC2 protein.  
XX  
KW Human; MMSC2; MMAC1; PDZ domain; tumour suppressor; tyrosine phosphatase;  
KW scaffolding protein; cancer.  
XX  
OS Homo sapiens.  
XX  
PN WO958548-A1.  
XX  
PD 18-NOV-1999.  
XX  
PE 07-MAY-1999; 99MO-US09969.  
XX  
PR 08-MAY-1998; 98US-0084740.  
XX  
PA (MYRI-) MYRIAD GENETICS INC.  
XX  
PI Bartel PL, Tavtigian SV;



xx MPI: 2000-053077/04.  
 DR N-PSDB: AA236453.  
 xx  
 PT Nucleic acids and polypeptides representing human MMS2, useful for  
 detecting, diagnosing a predisposition to, and treating cancer -  
 xx  
 PS Claim 1, Page 93-99; 112pp: English.  
 CC The present sequence represents human MMS2 protein. The MMS2 protein  
 CC binds to MMS2. The MMS2 protein has 11 post-synaptic density protein,  
 CC disc-large, 20-1 (PDZ) domains and one or more of these domains  
 CC interacts specifically with the carboxyl terminal amino acids of MMS2  
 CC (see AY53754). Specifically, it appears that domain 7, 10 and 13  
 CC interact with MMS2. Since MMS2 contains 11 PDZ domains and interacts  
 CC with MMS2, a known tumour suppressor having a region of homology with  
 CC protein tyrosine phosphatases, MMS2 acts as a scaffolding protein in a  
 CC common biological pathway with MMS2. It is believed that the  
 CC interaction between MMS2 and MMS2 is required for the tumour suppressor  
 CC activity of MMS2. The MMS2 polypeptides, polynucleotides, fragments and  
 CC specific or complex specific antibodies may be used for detecting cancer  
 CC or a predisposition to cancer and screening for agents that may be used  
 CC to treat MMS2 and/or MMS2 related cancer. The polypeptides and  
 CC polynucleotides may also be used to treat cancer.  
 SO Sequence 2037 AA:  
 Query Match 98.9%; Score 6914.5; DB 21; Length 2037;  
 Best Local Similarity 96.7%; Pred. No. 0;  
 Matches 1363; Conservative 7; Mismatches 3; Indels 37; Gaps 1;  
 QY 1 MYCCRRTPPTQSELSLDCIDIELEKPHYDLGEFTGSSSTEDPVLTAMTADGSTEY 60  
 DB 628 MGCCTTTPPTQSELSLDCIDIELEKPHYDLGEFTGSSSTEDPVLTAMTADGSTEY 687  
 QY 61 QAPLAWMEAGIOHIELEKSGKGLGFSIYDPIIDPASTVIRIIRSLVPGIAGKGRLLP 120  
 DB 688 GQPLAWMEAGIHLIELEKSGKGLGFSIYDPIIDPASTVIRIIRSLVPGIAGKGRLLP 747  
 QY 121 GDRLEFVNDVNLNENSLLEAVALKGAPSGYTRIGVAKRLPSPEEGYSAKEDSFLRP 180  
 DB 748 GDRLEFVNDVNLNENSLLEAVALKGAPSGYTRIGVAKRLPSPEEGYSAKEDSFLRP 807  
 QY 181 HSCERAGLADKPLFRADLALVTGNADLVDESTFSPSPENDSYSTQASTLSHGSSC 240  
 DB 808 HSCERAGLADKPLFRADLALVTGNADLVDESTFSPSPENDSYSTQASTLSHGSSC 867  
 QY 241 GDGLNYGSSLPSPPKVDIENSCDPLDLHMSLEELYTONLERODENTPVDISGSPAS 300  
 DB 868 GDGLNYGSSLPSPPKVDIENSCDPLDLHMSLEELYTONLERODENTPVDISGSPAS 927  
 QY 301 GTTINDYPPANAIEQOYECEENTVWTEHSLPSEVTSAPLEPSVLPDSAGSGEHLLEQSS 360  
 DB 928 GTTINDYPPANAIEQOYECEENTVWTEHSLPSEVTSAPLEPSVLPDSAGSGEHLLEQSS 987  
 QY 361 LACNAECVLAQVNSKESFERTINIAKGNSSLTMTVSANKDGMVRSIRHGAISROGR 420  
 DB 988 LACNAECVLAQVNSKESFERTINIAKGNSSLTMTVSANKDGMVRSIRHGAISROGR 1047  
 QY 421 IAIQDCILISINEESTISVTNAQARALRRHSLIGPDIKITYPAHLEEFKISLQOOSGR 480  
 DB 1048 IAIQDCILISINEESTISVTNAQARALRRHSLIGPDIKITYPAHLEEFKISLQOOSGR 1107  
 QY 481 WVALDIFSSYTRGROIPEREPREGESESELONTAYSNMORRRVRLMEPEKSGISIV 540  
 DB 1108 WVALDIFSSYTRGROIPEREPREGESESELONTAYSNMORRRVRLMEPEKSGISIV 1167  
 QY 541 GGRGMSRLSNCEVNRGIFIKHVLSDSPAGKNGTLKPGDRIYE----- 583  
 DB 1168 GGRGMSRLSNCEVNRGIFIKHVLSDSPAGKNGTLKPGDRIYE----- 1227  
 QY 584 -----APSQSSSEPEKAPLCSVPPPPSAPAEWGSMDHTOSSASKI 623

DB 1228 IAKAGPVVFMVQSLINPRPSGSESEPEKAPLCSVPPPPSAPAEWGSMDHTOSSASKI 1287  
 QY 624 SODVDKEDDEFGYSWKNIERRGTITGELHMIIELEKSGSLGSLAGNKRDRSMVFIYGI 683  
 DB 1288 SQDVKEDDEFGYSWKNIERRGTITGELHMIIELEKSGSLGSLAGNKRDRSMVFIYGI 1347  
 QY 684 DPNKAAGKGRQIADDELLETNGOILYGRSHONASSITKCAPSKYKIFIRKADVNOMA 743  
 DB 1348 DPNKAAGKGRQIADDELLETNGOILYGRSHONASSITKCAPSKYKIFIRKADVNOMA 1407  
 QY 744 VCPGNAVEPLPSNSENLOKTEPEPTVTSDAAVDLSSEFKMVOHLELPKDOGLGIAISEE 803  
 DB 1408 VCPGNAVEPLPSNSENLOKTEPEPTVTSDAAVDLSSEFKMVOHLELPKDOGLGIAISEE 1467  
 QY 804 DTLSCVITIKSTHEGVAAATDGRKAGDOILAVDDEIVYGVPIEKTSILKTAAMTVKLT 863  
 DB 1468 DTLSCVITIKSTHEGVAAATDGRKAGDOILAVDDEIVYGVPIEKTSILKTAAMTVKLT 1527  
 QY 864 HAENDSOAVPSAAGAAGEKNSOSIMVPOGSPPESTIRNTSSRPAFAADPATC 923  
 DB 1528 HAENDSOAVPSAAGAAGEKNSOSIMVPOGSPPESTIRNTSSRPAFAADPATC 1587  
 QY 924 PIIPCETIIEISKRTGIGSLIVGSDTLGAFIIEHYEEGAACKDGRMLAGDOILEV 983  
 DB 1588 PIIPCETIIEISKRTGIGSLIVGSDTLGAFIIEHYEEGAACKDGRMLAGDOILEV 1647  
 QY 984 NCIDLRKATHEAIVNROTPOVRRLTLRYDEAPRYKEEVCOTLTIELOKKFGKGLST 1043  
 DB 1648 NCIDLRKATHEAIVNROTPOVRRLTLRYDEAPRYKEEVCOTLTIELOKKFGKGLST 1707  
 QY 1044 VKRRNDGVGFVSDIYKGGIADPDGRLODOITLVNGEDVRNASEDAVAALIKCSLGYVT 1103  
 DB 1708 VKRRNDGVGFVSDIYKGGIADPDGRLODOITLVNGEDVRNASEDAVAALIKCSLGYVT 1767  
 QY 1104 LEVGRIKAGPFRSEERRPSQTSQVSEGLSSFTPLSGSSSTESSSSKKNALASEIOGL 1163  
 DB 1768 LEVGRIKAGPFRSEERRPSQTSQVSEGLSSFTPLSGSSSTESSSSKKNALASEIOGL 1827  
 QY 1164 RTVEKKKGPSTDSLGSINGVSSPLGDVPIFAMNHPPGVAAQOTKRLRGDRIYTCGS 1223  
 DB 1828 RTVEKKKGPSTDSLGSINGVSSPLGDVPIFAMNHPPGVAAQOTKRLRGDRIYTCGS 1887  
 QY 1224 TEGMHTQAVNLKKNASGSIEMQVAVAGDVSVTGHHOEPASSSLFTGLTSTIFODL 1283  
 DB 1888 TEGMHTQAVNLKKNASGSIEMQVAVAGDVSVTGHHOEPASSSLFTGLTSTIFODL 1947  
 QY 1284 GPPCKSTLERGPDGLGFSIVGSGSPHGDLPITYKTVFAKGAASEDGRKLRGDOITAV 1343  
 DB 1948 GPPCKSTLERGPDGLGFSIVGSGSPHGDLPITYKTVFAKGAASEDGRKLRGDOITAV 2007  
 QY 1344 NGOSLEGVTHEEAVAILKRTKGTVTLMVLS 1373  
 DB 2008 NGOSLEGVTHEEAVAILKRTKGTVTLMVLS 2037  
 RESULT 5  
 ID AAY04731 standard; Protein; 1005 AA.  
 AC AAY04731;  
 DT 06-JUL-1999 (first entry)  
 DE Mature protein containing PDZ domain from clone 38-2-1.  
 XX PDZ domain; gene expression; human umbilical vascular endothelial cell;  
 KM HUVEC; stimulation; tumour necrosis factor; TNF; protein binding;  
 XX cell; proliferation disorder; cancer.  
 OS Homo sapiens.  
 XX WO9907846-A1.  
 PN

XX 18-FEB-1999.  
 PD 12-AUG-1998; 98WO-JP03603.  
 XX 19-JUN-1998; 98JP-0189944.  
 PR 12-AUG-1997; 97JP-0230356.  
 XX (CHUG-) CHUGAI RES INST MOLECULAR MEDICINE INC.  
 PA Funahashi S, Miyata S;  
 PI WPI; 1999-167423/14.  
 DR N-PSDB; AAX29908.  
 XX  
 PT Protein containing PDZ domain, whose expression is enhanced by TNF  
 PT stimulation - plays an important role in protein/protein  
 PT interactions and is used for screening for proteins for use in  
 PT treatment of cell proliferation disorders such as cancer  
 PS  
 PS Claim 1; Page 84-87; 240pp; Japanese.  
 CC This sequence represents the mature portion of a new protein containing  
 CC a PDZ domain encoded by the clone 38-2-1, whose expression in human  
 CC umbilical vascular endothelial cells (HUVEC) is enhanced by stimulation  
 CC with tumour necrosis factor (TNF) alpha. The new protein is used to  
 CC identify proteins which bind to it (particularly to the PDZ domains) and  
 CC the genes encoding them, for use in the treatment of cell proliferation  
 CC disorders such as cancer.  
 CC  
 XX Sequence 1005 AA;  
 SQ  
 Query Match 72.7%; Score 5085; DB 20; Length 1005;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1005; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 369 MLQVSKSEFRTINIAKGNSSLGMTVSANKDGLMTVRSIIHGSAISRGRGAIAGCIL 428  
 DB 1 mlgnvskseftrtlniakgnsslgmtvsankdglmtvrsiihgalsrdgraiagcill 60  
 QY 429 SINESTISVTNAQARMLRHSILIGPDIKITYPAEHLFEKSLIGQOSRWALDIFS 488  
 DB 61 sineestisvtvnaqarmlrhtslilgpdikitypaehlfefkslsgqsrwalidifs 120  
 QY 489 SYTGRIPELPEREGGESEELONTAYSMNQPRVLELWREPSKSLGISTVGRGMSR 548  
 DB 121 sytgripelperegegeeseldntaysnmngrrvelwrepkslsgistvgrgmgr 180  
 QY 549 LSNGEVMRGIFIKVNLVLEDSPPAGKNGITAKPGDRIVEAPSSQSEPEKAPLCSVPPPSAF 608  
 DB 181 lsngevmrgrfikvnlvledspagkngitlkpgdriveapssqsepekaplcsvpppsaf 240  
 QY 609 AEMSDHTOSSASKISODVDKEDFEFGYSMKNIRRYGTLGELIMITELEKSHSLGSLA 668  
 DB 241 aemsdhtossaskisodvdkedeefegysmknierrygtlgeilmiteleyshslgslsla 300  
 QY 669 GNRKDSRMASVFIIVGIDPNGAAGKDRLOIADDELLEINGOILYGRSHONASSIICAFSKV 728  
 DB 301 gnrkdsrmavfiivgidpngaagkdrloiaddelleingqilysrshnaslkcapskv 360  
 QY 729 KIIRIRKDAVNQNAVCPGNAVEPLPSNSENLOKREPEVYTSDDAVALDSSEKNOHLE 788  
 DB 361 kiirirkdavnqnavcpgnaveplpsnsenlqketepevttsdaavdlssfkvnghle 420  
 QY 789 LPRKOGGLGTAISEEDTLGVIIRKSLTEHGVAAADGRLKXGDOILAADDDEVYVPIEKF 848  
 DB 421 lprkoggglgtaiseedtlsgviiirksltehgvaaadgrlkxvgdqlaavdddevyvpiekf 480  
 QY 849 ISLTKTKMKTVKLTIIHAENPDSQAVPSAAGAASEKKNSSQSLWPVPGSPPEPSINTS 908  
 DB 481 isltktkmkvtvklthienpdsqavpsaagaasekknssqslwvpgsppepsintts 540

QY 909 RSTPAIFASDPATPIIPGCEETIETISKRTGSLGSLVSGSDTLGAFIIEVEEGAA 968  
 DB 541 rsttpaifasdpatcpiipgceetietskrtgslgslvsgsdtlgafiiheveegaa 600  
 QY 969 CKDGRINAGDOILEVNGIDLRKATHDEAINVLRQTPORVRLTLRDEAPYKEEVCOTLT 1028  
 DB 601 ckdgrlwnagddilevngidlrkatthdeainvrlrtpqrvrltlrdeapykeevcdlt 660  
 QY 1029 TELKKKKGKGLGSLVSKRNDTGVEVDIVKGGIADDDGRILGDDOLLVNGEDVRNASQ 1088  
 DB 661 telkkkpgkglgslvskrntdgvfsvdivkggiaaddgrilgddollvngedvrnasq 720  
 QY 1089 EAVAALLKCSLGTVTLVGRKACPFHSERRPSQTSQVSEGLSFFFPPLSGSSTSESL 1148  
 DB 721 eavaallkcslgtvtlvgrkacpfhserrpsqtsqvseslsfffpplsastsesle 780  
 QY 1149 SSKKNALASFTGLRVEVKKGPTDLSGISIAGVSPGLADVFIFLMMHPGVAAQTO 1208  
 DB 781 sskknalaesftglrvemkkgptdlsigislagvsgpladvpiflammhpgvvaqtq 840  
 QY 1209 KLRYGDRIVTICGSTEGMTHTQAVNLKKNASGSIEMQVAGGVSVYTGHHOBPASSL 1268  
 DB 841 klrygdrivtfcgstegmthtqavnllknasgsiemqvaggvsvytgghobpassl 900  
 QY 1269 SFTGLTSTSIFFDDLGPPCKSITLERGPDLGFSIVGSGSPHGLPIYKTVFARKAA 1328  
 DB 901 sftglststsiifddlgppckstlergpdlgfsivsgspghlpiykvtfarkaa 960  
 QY 1329 SEDGRILKRGDIIIVNGOSLEGVTHEEVAVALIKRTKTVTLMLVLS 1373  
 DB 961 sedgrilkrdqilavngslegvtheevaualikrtkvtvlmlvls 1005

## RESULT 6

AAB01383  
 ID AAB01383 standard; Protein; 856 AA.

AC AAB01383;  
 DT 20-OCT-2000 (first entry)  
 XX

DE Neuron-associated protein.

XX Neuron associated protein; NEUAP; neurological disorder; epilepsy;  
 KW ischemic cerebrovascular disease; stroke; cerebral neoplasm;  
 KW Alzheimer's disease; Pick's disease; Huntington's disease;  
 KW dementia; Parkinson's disease; demyelinating disease; meningitis;  
 KW prion disease; kuru; Creutzfeldt-Jakob disease; neurofibromatosis;  
 KW cerebral palsy; muscular dystrophy; central nervous system; CNS;  
 KW peripheral nervous system; PWS; myopathy; schizophrenia;  
 KW actinic keratosis; arteriosclerosis; atherosclerosis; burns;  
 KW cirrhosis; hepatitis; mixed connective tissue disease; MCMD;  
 KW myelofibrosis; paroxysmal nocturnal haemoglobinuria; cancer;  
 KW autoimmune disease; inflammation; acquired immunodeficiency syndrome;  
 KW AIDS; Addison's disease; adult respiratory distress syndrome;  
 KW allergy; ankylosing spondylitis; amyloidosis; anaemia; asthma;  
 KW Werner syndrome; trauma; human.

OS Homo sapiens.  
 XX

FH Key Location/Qualifiers  
 FT Domain 1..65  
 FT Modified-site /label= PDZ domain signature 18  
 FT Modified-site /note= "Potential glycosylation site" 20  
 FT Modified-site /note= "Potential phosphorylation site" 69  
 FT Modified-site /note= "Potential phosphorylation site" 71  
 FT Modified-site /note= "Potential phosphorylation site" 73  
 FT Modified-site /note= "Potential phosphorylation site"

FT Modified-site 118 /note- "Potential phosphorylation site"  
 FT Modified-site 119 /note- "Potential phosphorylation site"  
 FT Domain 136..218 /label- PDZ domain signature  
 FT Modified-site 144..147 /note- "Glycosaminoglycan attachment site"  
 FT Modified-site 199 /note- "Potential glycosylation site"  
 FT Modified-site 238 /note- "Potential phosphorylation site"  
 FT Modified-site 253 /note- "Potential phosphorylation site"  
 FT Modified-site 263 /note- "Potential phosphorylation site"  
 FT Domain 269..349 /label- PDZ domain signature  
 FT Modified-site 284 /note- "Potential phosphorylation site"  
 FT Modified-site 296 /note- "Potential phosphorylation site"  
 FT Modified-site 337 /note- "Potential phosphorylation site"  
 FT Modified-site 341 /note- "Potential phosphorylation site"  
 FT Modified-site 369 /note- "Potential glycosylation site"  
 FT Modified-site 370 /note- "Potential phosphorylation site"  
 FT Modified-site 386 /note- "Potential phosphorylation site"  
 FT Modified-site 389 /note- "Potential glycosylation site"  
 FT Modified-site 390 /note- "Potential phosphorylation site"  
 FT Modified-site 414 /note- "Potential phosphorylation site"  
 FT Domain 415..497 /label- PDZ domain signature  
 FT Modified-site 475 /note- "Potential phosphorylation site"  
 FT Domain 511..592 /label- PDZ domain signature  
 FT Modified-site 531 /note- "Potential glycosylation site"  
 FT Modified-site 568 /note- "Potential glycosylation site"  
 FT Modified-site 599 /note- "Potential phosphorylation site"  
 FT Modified-site 604 /note- "Potential phosphorylation site"  
 FT Modified-site 625 /note- "Potential phosphorylation site"  
 FT Modified-site 633 /note- "Potential phosphorylation site"  
 FT Modified-site 634 /note- "Potential phosphorylation site"  
 FT Domain 648..733 /label- PDZ domain signature  
 FT Modified-site 690 /note- "Potential phosphorylation site"  
 FT Modified-site 705 /note- "Potential phosphorylation site"  
 FT Modified-site 721 /note- "Potential glycosylation site"  
 FT Domain 773..856 /label- PDZ domain signature  
 FT Region 819..821 /label- Cell attachment sequence  
 FT Modified-site 835 /note- "Potential phosphorylation site"  
 XX

PN MO200034477-A2.  
 XX 15-JUN-2000.  
 XX  
 PF 10-DEC-1999; 99WO-US30408.  
 XX  
 PR 11-DEC-1998; 98US-0210083.  
 PR 11-DEC-1998; 98US-9123456.  
 PR 09-FEB-1999; 99US-0119365.  
 PR 16-MAR-1999; 99US-0124687.  
 XX  
 PA (INCYTE) INCYTE PHARM INC.  
 XX  
 PI Tang YT, Yue H, Baughn MR, Hillman JL, Lal P, Au-young J, Yang J;  
 PI Lu DAM, Azimzal Y;  
 XX  
 DR WPJ; 2000-423423/36.  
 DR N-PSDB; AAA47424.  
 XX  
 PT New human neuron-associated proteins and polynucleotides encoding them,  
 PT useful for diagnosis, treatment and prevention of cell proliferative  
 PT disorders including cancer, neuronal and neurological disorders  
 XX  
 PS Claim 1; Page 105-108; 145pp; English.  
 XX  
 CC Human neuron-associated proteins (NEUAP) can be used for  
 CC treating or preventing a disorder associated with decreased  
 CC expression or activity of NEUAP. Antagonists of NEUAP are useful for  
 CC treating or preventing disorder associated with increased expression  
 CC or activity of NEUAP. NEUAP or their fragments or derivatives are  
 CC useful for treating neurological disorder such as epilepsy, ischemic  
 CC cerebrovascular disease, stroke, cerebral neoplasms, Alzheimer's  
 CC disease, Pick's disease, Huntington's disease, dementia and  
 CC Parkinson's disease. NEUAPs are also useful for treating other  
 CC demyelinating diseases, bacterial and viral meningitis, prion  
 CC diseases including kuru, Creutzfeldt-Jakob disease, nutritional and  
 CC metabolic diseases of the nervous system, neurofibromatosis, other  
 CC developmental disorders of the central nervous system, cerebral  
 CC palsy, neuroskeletal disorders, autonomic nervous system disorders,  
 CC cranial nerve disorders, spinal cord diseases, muscular dystrophy and  
 CC other neuromuscular disorders, peripheral nervous system disorders,  
 CC inherited, metabolic, endocrine, and toxic myopathies, mental  
 CC disorders including mood, anxiety and schizophrenic disorders, a cell  
 CC proliferative disorder such as actinic keratosis, arteriosclerosis,  
 CC atherosclerosis, bursitis, cirrhosis, hepatitis, mixed connective  
 CC tissue disease (MCTP), myelofibrosis, paroxysmal nocturnal  
 CC haemoglobinuria, cancers of the adrenal gland, bladder, bone,  
 CC bone marrow, brain, breast, cervix, and an autoimmune/inflammatory  
 CC disorder such as acquired immunodeficiency syndrome (AIDS), Addison's  
 CC disease, adult respiratory distress syndrome, allergies, ankylosing  
 CC spondylitis, amyloidosis, anemia, asthma, Werner syndrome,  
 CC complications of cancer, hemodialysis, and extracorporeal circulation,  
 CC viral, bacterial, fungal parasitic, protozoal, and helminthic  
 CC infections, and trauma. This protein was given the Incyte ID no.  
 CC 3201753CD1.  
 XX  
 SO Sequence 856 AA:  
 Query Match 58.4%; Score 4080.5; DB 21; Length 856;  
 Best Local Similarity 94.7%; Pred. No. 0;  
 Matches 811; Conservative 5; Mismatches 3; Indels 37; Gaps 1;  
 Oy 555 MKGIFKHVLEDPAGKNGTKPGDRIVE----- 583  
 Db 1 mrglfkhvleddspgkngtklpgdrivevdmldrdashegaveairkagnpvfwqgs 60  
 Oy 584 -----APSQSESEPERKPLCSVPPPPSAFAEKSDHTOSASAKTISDDVKKEDPFQGSW 637  
 Db 61 llnrpapsqsepekaplcsvppppsaftbmsdhtqsaaslsqdvkdeafgysw 120  
 Oy 638 KNIRRYGTLNGELMILKRGHSGLGSLAGNKRDRMSVFIVGIDNGAAGKGRLOI 697  
 |||



Db 601 rvgddivlvcstegmthtgavnlknasgsiemvavagdvsvltghhgpassisf 660  
 Qy 1271 TGLTSTSTFODDLGPPOCKSTLREPGDLGPIVGYGSPHGDPIYKTYFAKGAASE 1330  
 Db 661 tglststfoddlgppgckstlrepgdlgpiygyyspghdpliyktyfakgaase 720  
 Qy 1331 DGRLRGDDIIVANGOSLEGVTHEEVAAILKRTKGTVTLMLVLS 1373  
 Db 721 dgrlkrqgdlavngsglegvtheeavallkrtkgtvltlmvls 763

## RESULT 8

AA04734  
 ID AAY04734 standard; Protein: 1239 AA.

AC AAY04734;  
 DT 06-JUL-1999 (first entry)  
 DE Protein containing PDZ domain from clone 38-2-1c.  
 XX  
 XX PDZ domain; gene expression: human umbilical vascular endothelial cell;  
 KM HUVEC; stimulation: tumour necrosis factor; TNF; protein binding;  
 KM cell; proliferation disorder; cancer.  
 XX  
 OS Homo sapiens.

XX MO9907846-A1.  
 XX  
 PD 18-FEB-1999.  
 XX  
 PF 12-AUG-1998; 98WO-JP03603.  
 XX  
 PR 19-JUN-1998; 98JP-0189944.  
 PR 12-AUG-1997; 97JP-0230356.  
 XX  
 PA (CHUGAI) CHUGAI RES INST MOLECULAR MEDICINE INC.

XX Funahashi S, Miyata S;  
 PI WPI: 1999-167423/14.  
 DR N-PSDB; AAX29911.  
 XX  
 XX  
 PT Protein containing PDZ domain, whose expression is enhanced by TNF  
 PT stimulation - plays an important role in protein/protein  
 PT interactions and is used for screening for proteins for use in  
 PT treatment of cell proliferation disorders such as cancer  
 XX  
 PS Claim 1; Page 177-182; 240pp; Japanese.  
 XX  
 CC This sequence represents a new protein containing a PDZ domain encoded  
 CC by the clone 38-2-1c, whose expression in human umbilical vascular  
 CC endothelial cells (HUVEC) is enhanced by stimulation with tumour necrosis  
 CC factor (TNF) alpha. The new protein is used to identify proteins which  
 CC bind to it (particularly to the PDZ domains) and the genes encoding them,  
 CC for use in the treatment of cell proliferation disorders such as cancer.  
 XX  
 SO Sequence 1239 AA;

Query Match 42.6%; Score 2981; DB 20; Length 1239;  
 Best Local Similarity 99.7%; Pred. No. 1,1e-231;  
 Matches 581; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MYCCRRVPTPTQSELDLDLDELTEKPHVDLGEFTSSSETEDPVLAMDAGOSTERY 60  
 Db 628 myccrrvptptqseldldldeltekphvdlgeftsssepdpvlamdagosteey 687  
 Qy 61 GQPLAMMEGIOHIELEKSGKIGFSLIDYODIPASVITIRSRVPGIAGKGRLLP 120  
 Db 688 gqplammegiohieleksgkigfslidyodipasviltirsrvpigiaeckgrllp 747  
 Qy 121 GDRLMFVNDVNLNLSLEAEALKGAPSGTVRIGVAKPLPLSPGEGYSAKEDSFLYPP 180

Db 748 gdrlmfvndvnlensleaeaalkgapsgtvrlgvakplplspeggyvakedsflypp 807  
 Qy 181 HSCERAGLADKPLFRADIALVGTNDADLVDESTFSPSPENDSTYSTQASLSLHGSSC 240  
 Db 808 hsceragladkplfradialvgtndadvdestfespspendstysqasllshgssc 867  
 Qy 241 GGLNKGSSLPSPKVDIENSCDPLDLHMSLEELTYQNLLERODEMTSPVDSMGAS 300  
 Db 868 gglngksslpsspkdienscdpvldlhmsleelyqnllerodemtspvdsimgas 927  
 Qy 301 GFTINDYTPANAIEQOYECENTIVWTESHLPSEVISSEALPSLPDSAGKSGEHLLEQSS 360  
 Db 928 gftindytpanaieqoyecentivwteshlpsevissealpslpdsagksgehllegss 987  
 Qy 361 IACNAECVWLQNVSKESFERTINIAKNSLQMTVYSAKNDGLMTVRSIIHGASISDGR 420  
 Db 988 iacnaecvmlqnvskesfertinlakngslqmtvsankdglmvtvrsllhgaisrdgr 1047  
 Qy 421 IAIIGCIITINEESTISTVNMQARMLRRHSIGDITITYPAHLEEFKISLGOOSGR 480  
 Db 1048 iaigcilitineestistvnaqaramlrrhsllgpdiklyvpaehleefkisl9qsgsr 1107  
 Qy 481 VVALDIFSSYTGSDIPELPEREGEGESELQNTAYSNMNOPRRYELWREPSSLGISIV 540  
 Db 1108 vvaldifssytgdiplpereegegeeselnqntaysnmnoprryelwrepsslgisliv 1167  
 Qy 541 GGRGMSRLSNCEVMRGTFIKHVEDSPAGKNGTKPGDRIVE 583  
 Db 1168 ggrgmssrlsncevmrgtfikhvedspagkngtkpgdrive 1210

## RESULT 9

AA24025  
 ID AAY24025 standard; Protein: 1881 AA.

AC AAY24025;  
 DT 29-SEP-1999 (first entry)  
 XX  
 XX  
 DE Amino acid sequence of the human MMS1 protein.  
 XX  
 KM Human; MMS1 protein; MMS1 interacting protein; tumour suppression;  
 KM MMAC1 pathway; immunogen; cancer; cell neoplastic growth.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO9936566-A1.  
 XX  
 PD 22-JUL-1999.  
 XX  
 PF 19-JAN-1999; 99WO-US00095.  
 PR 20-JAN-1998; 98US-0071861.  
 XX  
 PA (MYRI-) MYRIAD GENETICS INC.

XX Bartel PL, Tavtigian SV;  
 PI WPI: 1999-458472/38.  
 DR N-PSDB; AAX86366.  
 XX  
 PT MMS1, an MMAC1 (tumour suppressor) interacting protein and related  
 PT polynucleotides  
 XX  
 PS Claim 14; Page 88-93; 107pp; English.  
 XX  
 CC The present sequence represents a MMS1 protein. The protein is a MMAC1  
 CC interacting protein which is involved in tumour suppression activity  
 CC in the MMAC1 pathway. MMS1, antigenic fragments or fusion proteins of  
 CC these are used as immunogens for antibody production. Primers derived  
 CC from MMS1 genomic clones can be used for identification of MMS1 genes  
 CC and for synthesis by amplification of MMS1 DNA or RNA. Detecting an



XX MPI: 1999-621386/54.  
 DR N-PSDB; AA252969.  
 XX  
 PT New human nucleic acid sequences from pancreatic tumors, and related  
 proteins  
 XX  
 PS Claim 23; Page 449-450; 502pp; German.  
 XX  
 CC This invention describes novel polypeptides and their encoding nucleic  
 CC acids derived from human pancreatic tumor tissue which have cytostatic  
 CC activity. The sequences are also useful in producing pharmaceutical  
 CC compositions for treatment of pancreatic tumors. AA73814-74252  
 CC represent protein fragments encoded by the human pancreatic tumor CDNA  
 CC library derived expressed sequence tag (EST) sequences represented in  
 CC AA52858-253014.  
 CC  
 XX Sequence 318 AA;  
 SQ

Query Match 21.5%; Score 1501; DB 20; Length 318;  
 Best Local Similarity 96.7%; Pred. No. 6e-113;  
 Matches 297; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

OY 790 PKDGGGLGIAISEEDTLSCVITKSLTEHGVATDGRKLVGDQILAVDDEIVVGYPIEKPI 849  
 DB 9 prrggslglaisedtlsgvltksltelhgvatdgrlkvgdqlavddeivvgypiekfi 68  
 OY 850 SLTKAKMTVKLTTHAENPDQAVPSAAGASGEEKNSSQSLMVPQSGPEPESIRNTSR 909  
 DB 69 sltkakmtvklthnaenpdqavpsaagaasekknssqslmvpqsgpepesirntsr 128  
 OY 910 SSTPAIFASDPATCPIPCCECTTETISKGRGSLSYVSGSTLLGAFTHHYVEEGAAC 969  
 DB 129 sstpaifasdpctpiipccettetlskgrtlysgstllgafthhyveegaac 188  
 OY 970 KDGRLMAGDQILEVNGIDLRKATHDEAINVLKQTFQRYRLTLRYDEAPYKEEVCDTLTI 1029  
 DB 169 kdgrlmagdqilevngidlrkatheainvltkqtforyrltlrydeapykeevcdtlti 248  
 OY 1030 ELQKPPGKGLSYVGRNDGVFVSDIVKGGIADPGRLLIGDQILLVNGEDVRNASE 1089  
 DB 249 elqkppkgglisyvgrndgvfvsdivkgyiadadgllmgdqlimvngedvrnatge 308  
 OY 1090 AVAALLK 1096  
 DB 309 avavwvk 315

RESULT 11  
 ID AA04738 standard; Protein; 251 AA.  
 XX  
 AC AA04738;  
 XX

06-JUL-1999 (first entry)

PDZ domain-containing protein gene encoded by clone FH750.

XX PDZ domain; gene expression; human umbilical vascular endothelial cell;  
 KW HUVEC; stimulation; tumour necrosis factor; TNF; protein binding; PCR;  
 KW cell; proliferation disorder; cancer; primer; amplification.  
 XX

OS Homo sapiens.

XX WO9907846-A1.

PN 18-FEB-1999.

PD 12-AUG-1998; 98WO-JP03603.

XX 19-JUN-1998; 98JP-0189944.

PR 12-AUG-1997; 97JP-0230356.

XX (CHUG-) CHUGAI RES INST MOLECULAR MEDICINE INC.  
 PA Funahashi S, Miyata S;  
 XX  
 PI MPI: 1999-167423/14.  
 DR N-PSDB; AA29969.  
 XX

PT Protein containing PDZ domain, whose expression is enhanced by TNF  
 PT stimulation - plays an important role in protein/protein  
 PT interactions and is used for screening for proteins for use in  
 PT treatment of cell proliferation disorders such as cancer  
 XX

PS Example 7; Page 149-151; 240pp; Japanese.

XX This sequence represents a new protein containing PDZ domains encoded  
 CC by clone FH750, whose expression in human umbilical vascular endothelial  
 CC cells (HUVEC) is enhanced by stimulation with tumour necrosis factor  
 CC (TNF) alpha. The new protein is used to identify proteins which bind  
 CC to it (particularly to the PDZ domains) and the genes encoding them,  
 CC for use in the treatment of cell proliferation disorders such as cancer.  
 CC  
 XX Sequence 251 AA;  
 SQ

Query Match 18.1%; Score 1266; DB 20; Length 251;  
 Best Local Similarity 99.6%; Pred. No. 4.1e-94;  
 Matches 250; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 341 PSVLPDSAGKSGEHLLEOSSLACNAECVMLQNVKSEFERTINAKGSSLGMTVSANKD 400  
 DB 1 psvlpdsagksgeyllsgslacnaecvmlqnvksesfertlnakgsslgmtvsankd 60  
 OY 401 GLGKTVRSIHGCAISRGRALIGDCIISINEESTISVTNMOARMLRHSILGPDITIT 460  
 DB 61 glgktvrsihgcaisrgraligdcilisineestisvtnmoarmlrshilgpditit 120  
 OY 461 YVPAHLEFFKSTLGOQGRVVALDIFSSYGRDIPERPEREGSGESLONTAYSMNN 520  
 DB 121 yvpaehleefkstlsgqgrvvaldifsygrdipeperegsgeslontaysmnn 180  
 OY 521 QPRVLEWRPESKSLGISVGRGMSRLSNGEVARGIFIKHVLDSPPAKNGTLKPGDR 580  
 DB 181 qprvlewrpksksgisvgrgmsrslsngevargifikhvledspakngtlkpgdr 240  
 OY 581 IVEAPSOESE 591  
 DB 241 iveapsgese 251

RESULT 12  
 ID AA04739 standard; Protein; 272 AA.  
 XX  
 AC AA04739;  
 XX

06-JUL-1999 (first entry)

PDZ domain-containing protein gene encoded by clone FH850.

XX PDZ domain; gene expression; human umbilical vascular endothelial cell;  
 KW HUVEC; stimulation; tumour necrosis factor; TNF; protein binding; PCR;  
 KW cell; proliferation disorder; cancer; primer; amplification.  
 XX

OS Homo sapiens.

XX WO9907846-A1.

PN 18-FEB-1999.

PD 12-AUG-1998; 98WO-JP03603.

XX 19-JUN-1998; 98JP-0189944.



PR 12-AUG-1997; 97JP-0230356.  
 XX (CHUG-) CHUGAI RES INST MOLECULAR MEDICINE INC.  
 XX  
 XX Funahashi S, Miyata S;  
 PI WPI: 1999-167423/14.  
 DR N-PSDB; AAX29970.  
 XX  
 XX Protein containing PDZ domain, whose expression is enhanced by TNF  
 PT stimulation - plays an important role in protein/protein  
 PT interactions and is used for screening for proteins for use in  
 PT treatment of cell proliferation disorders such as cancer  
 PS  
 PS Example 7; Page 152-154; 240pp; Japanese.  
 XX  
 XX This sequence represents a new protein containing PDZ domains encoded  
 CC by clone FH950, whose expression in human umbilical vascular endothelial  
 CC cells (HUVEC) is enhanced by stimulation with tumour necrosis factor  
 CC (TNF) alpha. The new protein is used to identify proteins which bind  
 CC to it (particularly to the PDZ domains) and the genes encoding them,  
 CC for use in the treatment of cell proliferation disorders such as cancer.  
 XX  
 SQ Sequence 272 AA;

Query Match 17.6%; Score 1228; DB 20; Length 272;  
 Best Local Similarity 99.6%; Pred. No. 5.6e-91;  
 Matches 242; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 Oy 341 PSVLPDSAGKSEHLLEQSSLAACVLMQNSKESFERTINIAKNSISGMTVSANKD 400  
 Db 1 psvlpdsagkseyllleqsslaacvmlqnskefertiflnaknsslgmtvsankd 60  
 Oy 401 GLGMIVRSIIHGCAISRDRRAIGDCILISTNEESTISVTNAQARMLRRHSLIGPDIKIT 460  
 Db 61 glgmivrsiihgaisrdgrlaigdcilistneestisvtnaqaramlrrhsligpdikit 120  
 Oy 461 YVPAHEHEEFKISIGQSGRYMALDIFSSYGRDIPELPEREEGEGESLQNTAVSNMN 520  
 Db 121 yvpaehheefkisygsgrymaldifssygrtdipelpereegegeselnqntavsnmn 180  
 Oy 521 QPRVELMREPSKSLGISIVGRMGSRSLNGEVMRGIPTIKHVLSDSPACKNGTLKPGDR 580  
 Db 181 qprvelmrepskslgisivgrmgssrlngewmrgiflikhvlstdspackngtlkpgdr 240  
 Oy 581 IVE 583  
 Db 241 lve 243

RESULT 13  
 AAY04740  
 ID AAY04740 standard; Protein: 319 AA.  
 XX  
 AC AAY04740;  
 XX  
 DT 06-JUL-1999 (first entry)  
 XX  
 DE PDZ domain-containing protein gene encoded by clone FH950.  
 XX  
 XX PDZ domain; gene expression; human umbilical vascular endothelial cell;  
 KW HUVEC; stimulation; tumour necrosis factor; TNF; protein binding; PCR;  
 KW cell; proliferation disorder; cancer; primer; amplification.  
 XX  
 OS Homo sapiens.  
 XX  
 XX PN WC907846-A1.  
 XX  
 PD 18-FEB-1999.  
 XX  
 PF 12-AUG-1998; 98MO-JP03603.  
 XX

PR 19-JUN-1998; 98JP-0189944.  
 XX 12-AUG-1997; 97JP-0230356.  
 XX  
 XX (CHUG-) CHUGAI RES INST MOLECULAR MEDICINE INC.  
 XX  
 XX Funahashi S, Miyata S;  
 PI WPI: 1999-167423/14.  
 DR N-PSDB; AAX29970.  
 XX  
 XX Protein containing PDZ domain, whose expression is enhanced by TNF  
 PT stimulation - plays an important role in protein/protein  
 PT interactions and is used for screening for proteins for use in  
 PT treatment of cell proliferation disorders such as cancer  
 PS  
 PS Example 7; Page 155-158; 240pp; Japanese.  
 XX  
 XX This sequence represents a new protein containing PDZ domains encoded  
 CC by clone FH950, whose expression in human umbilical vascular endothelial  
 CC cells (HUVEC) is enhanced by stimulation with tumour necrosis factor  
 CC (TNF) alpha. The new protein is used to identify proteins which bind  
 CC to it (particularly to the PDZ domains) and the genes encoding them,  
 CC for use in the treatment of cell proliferation disorders such as cancer.  
 XX  
 SQ Sequence 319 AA;

Query Match 17.6%; Score 1228; DB 20; Length 319;  
 Best Local Similarity 99.6%; Pred. No. 7.4e-91;  
 Matches 242; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 Oy 341 PSVLPDSAGKSEHLLEQSSLAACVLMQNSKESFERTINIAKNSISGMTVSANKD 400  
 Db 1 psvlpdsagkseyllleqsslaacvmlqnskefertiflnaknsslgmtvsankd 60  
 Oy 401 GLGMIVRSIIHGCAISRDRRAIGDCILISTNEESTISVTNAQARMLRRHSLIGPDIKIT 460  
 Db 61 glgmivrsiihgaisrdgrlaigdcilistneestisvtnaqaramlrrhsligpdikit 120  
 Oy 461 YVPAHEHEEFKISIGQSGRYMALDIFSSYGRDIPELPEREEGEGESLQNTAVSNMN 520  
 Db 121 yvpaehheefkisygsgrymaldifssygrtdipelpereegegeselnqntavsnmn 180  
 Oy 521 QPRVELMREPSKSLGISIVGRMGSRSLNGEVMRGIPTIKHVLSDSPACKNGTLKPGDR 580  
 Db 181 qprvelmrepskslgisivgrmgssrlngewmrgiflikhvlstdspackngtlkpgdr 240  
 Oy 581 IVE 583  
 Db 241 lve 243

RESULT 14  
 AAY74151  
 ID AAY74151 standard; Protein: 206 AA.  
 XX  
 AC AAY74151;  
 XX  
 DT 14-MAR-2000 (first entry)  
 XX  
 DE Human prostate tumor EST fragment derived protein #338.  
 XX  
 XX Pancreas; tumor; EST; expressed sequence tag; human; cytostatic;  
 KW treatment.  
 XX  
 OS Homo sapiens.  
 XX  
 XX PN DE19820190-A1.  
 XX  
 PD 04-NOV-1999.  
 XX  
 PF 28-APR-1998; 98DE-1020190.  
 XX

PR 28-APR-1998; 98DE-1020190.  
 XX (META-) METAGEN GES GENOMFORSCHUNG MBH.  
 PA Rosenthal A, Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E;  
 XX WPI; 1999-621386/54.  
 XX N-PSDB; AA52969.  
 DR New human nucleic acid sequences from pancreatic tumors, and related  
 XX proteins -  
 PT Claim 23; Page 450; 502pp; German.  
 XX  
 CC This invention describes novel polypeptides and their encoding nucleic  
 CC acids derived from human pancreatic tumor tissue which have cytostatic  
 CC activity. The sequences are also useful in producing pharmaceutical  
 CC compositions for treatment of pancreatic tumors. AA73814-Y74252  
 CC represent protein fragments encoded by the human pancreatic tumor cDNA  
 CC library derived expressed sequence tag (EST) sequences represented in  
 CC AA52858-253014.  
 XX  
 SQ Sequence 206 AA:

Query Match 13.3%; Score 928; DB 20; Length 206;  
 Best Local Similarity 95.3%; Pred. No. 6.5e-67;  
 Matches 184; Conservative 1; Mismatches 6; Indels 2; Gaps 1;

OY 606 SAFAEKSDHTOSSAKTSIQDYDEKDEFGYSWKNTIRERYGLTGLHMLEKSHSGIGL 665  
 DB 1 sfaeagschhtgsasaklsqdvkedefgyswknlrerygltlgtelhmlelksngsl 60  
 OY 666 SLAGKKDSKMSVFTVIGDIPNGAGKGRLOJADELEINQILYGRSHQNSSTIKCAP 725  
 DB 61 slagkkdsrmvflvyldpngaagkdgrlqadelldlengqillygrshqnsstikcap 120  
 OY 726 SKVKIIFIRNKDAVNOMAVCPGNAVEPLPSNSENLONKTEPTVTSDAAVLLSFKNVQ 785  
 DB 121 skvkliflrnkdaavnmaavcpgnaveplpsnsenlqnketpvttsdaavllsfknvq 180  
 OY 786 HLELPKDOGIGI 798  
 DB 181 h--sgasgsggrgv 191

## RESULT 15

AAV04736  
 ID AAV04736 standard; Protein; 590 AA.

XX  
 AC AAV04736;

DT 06-JUL-1999 (first entry)

DE PDZ domain-containing protein gene encoded by clone D-2.

XX PDZ domain: gene expression: human umbilical vascular endothelial cell;  
 KM HUVEC; stimulation; tumor necrosis factor; TNF; protein binding; PCR;  
 KM cell; proliferation disorder; cancer; primer; amplification.

OS Homo sapiens.

XX W09907846-A1.

PD 18-FEB-1999.

PF 12-AUG-1998; 98WO-JP03603.

PR 19-JUN-1998; 98JP-0189944.

XX 12-AUG-1997; 97JP-0230356.

(CHUG-) CHUGAI RES INST MOLECULAR MEDICINE INC.

PI Funahashi S, Miyata S;  
 XX WPI; 1999-167423/14.  
 DR N-PSDB; AAX29967.  
 XX  
 PT Protein containing PDZ domain, whose expression is enhanced by TNF  
 PT stimulation - plays an important role in protein/protein  
 PT interactions and is used for screening for proteins for use in  
 PT treatment of cell proliferation disorders such as cancer  
 XX  
 XX Example 7; Page 140-146; 240pp; Japanese.  
 CC This sequence represents a new protein containing PDZ domains encoded  
 CC by clone D-2, whose expression in human umbilical vascular endothelial  
 CC cells (HUVEC) is enhanced by stimulation with tumor necrosis factor  
 CC (TNF) alpha. The new protein is used to identify proteins which bind  
 CC to it (particularly to the PDZ domains) and the genes encoding them,  
 CC for use in the treatment of cell proliferation disorders such as cancer.  
 XX  
 SQ Sequence 590 AA:

Query Match 11.6%; Score 814; DB 20; Length 590;  
 Best Local Similarity 98.8%; Pred. No. 6.6e-57;  
 Matches 160; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 MYCCRRTPPTTQSLDLCIDITLTKPHVDLGEFGSSSETPVLAMDAGSTEEV 60  
 DB 429 myccrrtvppttqslldlciditltekphvdlgefgsssepepvlamtdagsteev 488  
 OY 61 GAPLAMEAGIOHILEKSGKGFSLDYODPIPASTVIRRLVGGIAEKGRLLP 120  
 DB 489 gaplameagiqhmlteksgkyglfslldyqdpdpastvllrlvpgglaekgrllp 548  
 OY 121 GDRLMFVNDVNLNSLSLEAAVEALKGAPSGTVRIGVAPRLP 162  
 DB 549 gdrlmfvndvnlensleaavealkgapsgtvrlgvakppl 590

Search completed: July 12, 2001, 14:40:26  
 Job time: 189 sec

Fri Jul 13 15:00:06 2001

us-09-502-698-1.rag

Page 14

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2001, 14:38:32 ; Search time 40.62 seconds  
(without alignments)  
2574.782 Million cell updates/sec

Title: US-09-502-698-1

Perfect score: 6990

Sequence: 1 MWCRRVPTTQSELDSD.....EENVAIIKRTKGVTLNVL 1373

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR1:\*  
2: PIR2:\*  
3: PIR3:\*  
4: PIR4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description                |
|------------|-------|-------------|--------|-------|----------------------------|
| 1          | 5811  | 83.1        | 2054   | 2     | T46612 multi PDZ domain p  |
| 2          | 5617  | 80.4        | 2055   | 2     | T30259 multiple PDZ domain |
| 3          | 962.5 | 13.8        | 2172   | 2     | T20145 hypothetical prote  |
| 4          | 589   | 8.4         | 728    | 2     | T09457 numb-binding prote  |
| 5          | 586   | 8.4         | 628    | 2     | T09458 numb-binding prote  |
| 6          | 534   | 7.6         | 1012   | 2     | T23160 hypothetical prote  |
| 7          | 523   | 7.5         | 2450   | 2     | S71625 protein-tyrosine-p  |
| 8          | 510.5 | 7.3         | 2294   | 2     | I67630 protein tyrosine p  |
| 9          | 509.5 | 7.3         | 2466   | 2     | I67629 protein tyrosine p  |
| 10         | 494.5 | 7.1         | 2490   | 1     | A54971 protein-tyrosine-p  |
| 11         | 453.5 | 6.5         | 1256   | 2     | JEO209 brain-specific ang  |
| 12         | 407.5 | 5.8         | 852    | 2     | T10811 channel associated  |
| 13         | 399.5 | 5.7         | 1131   | 2     | T15617 hypothetical prote  |
| 14         | 398.5 | 5.7         | 1277   | 2     | T14152 synaptic scaffold   |
| 15         | 397.5 | 5.7         | 870    | 2     | G01974 channel associated  |
| 16         | 385   | 5.5         | 1171   | 2     | T42372 probable guanylate  |
| 17         | 384   | 5.5         | 767    | 2     | T09599 postsynaptic densi  |
| 18         | 383   | 5.5         | 724    | 2     | JH0800 synapse-associated  |
| 19         | 380   | 5.4         | 720    | 2     | A45436 synapse-associated  |
| 20         | 379   | 5.4         | 960    | 1     | A38651 discs-large tumor   |
| 21         | 378.5 | 5.4         | 911    | 2     | I56552 synapse-associated  |
| 22         | 377   | 5.4         | 904    | 2     | I38757 homolog of Drosoph  |
| 23         | 377   | 5.4         | 926    | 2     | I38756 homolog of Drosoph  |
| 24         | 357   | 5.1         | 1112   | 2     | T33733 AMPA glutamate rec  |
| 25         | 353   | 5.1         | 1464   | 2     | T13711 bazooka gene prote  |
| 26         | 352   | 5.0         | 1337   | 2     | T13948 acylpeptide protein |
| 27         | 264   | 3.8         | 1360   | 2     | T34302 cell polarity prot  |
| 28         | 246   | 3.5         | 1163   | 2     | JEO366 tight junction pro  |
| 29         | 243   | 3.5         | 1736   | 2     | A47747 tight junction pro  |

|    |       |     |       |   |                           |
|----|-------|-----|-------|---|---------------------------|
| 30 | 237.5 | 3.4 | 1745  | 2 | A46431 tight junction-ass |
| 31 | 237   | 3.4 | 1495  | 2 | T31434 desmin-180 - rat   |
| 32 | 233.5 | 3.3 | 1281  | 2 | T00346 hypothetical prote |
| 33 | 228   | 3.3 | 1116  | 2 | I54378 gene x104 protein  |
| 34 | 216   | 3.1 | 87    | 2 | S60315 PSD-95-related pro |
| 35 | 214.5 | 3.1 | 1095  | 2 | T43275 neurabin - rat     |
| 36 | 210   | 3.0 | 1367  | 2 | T13703 tana protein - fru |
| 37 | 203.5 | 2.9 | 2232  | 2 | T34434 hypothetical prote |
| 38 | 202   | 2.9 | 13055 | 2 | T16580 hypothetical prote |
| 39 | 197.5 | 2.8 | 3562  | 2 | A47171 chondroitin sulfat |
| 40 | 197   | 2.8 | 505   | 2 | S62894 alpha-syntrophin - |
| 41 | 194.5 | 2.8 | 126   | 2 | T81210 tyrosine phosphata |
| 42 | 194   | 2.8 | 1829  | 2 | T24583 hypothetical prote |
| 43 | 191   | 2.7 | 723   | 2 | T14765 hypothetical prote |
| 44 | 190.5 | 2.7 | 1829  | 2 | T41751 1-afadin - rat     |
| 45 | 190   | 2.7 | 817   | 2 | T03852 protein phosphatas |

#### ALIGNMENTS

RESULT 1

T46612

multi PDZ domain protein 1 - rat

C:Species: Rattus norvegicus (Norway rat)

C:Date: 18-Feb-2000 #sequence: revision 18-Feb-2000 #text\_change 21-Jul-2000

C:Accession: T46612

R:Ullmer, C.; Schumack, K.; Figge, A.; Lubbert, H.

FEBS Lett. 424, 63-68, 1998

A:Title: Cloning and characterization of MUPP1, a novel PDZ domain protein.

A:Reference number: 223104; MUID:98196865

A:Accession: T46612

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-2054 <UL>

A:Cross-references: EMBL:AJ001320; NID:g2959978; PIDN:CAA04681.1; PID:g2959979

A:Experimental source: brain

C:Genetics:

A:Gene: MUPP1

| Query Match           | 83.1%            | Score 5811   | DB 2      | Length 2054 |
|-----------------------|------------------|--|-----------|-------------|
| Best local similarity | 79.5%            | Pred. No. 0  |           |             |
| Matches 1147          | Conservative 102 | Mismatches 116   | Indels 78 | Gaps 5      |
| QY                    | 1                | MWCRRVPTTQSELDSDLDLDIELEKRPVLDGEFISSETEDEPVLAMTAGOSTEEV    | 60        |             |
| DB                    | 620              | MWCRRVPTTQSELDSDLDLDIELEKRPVLDGEFISSETEDEPVLAMTAGOSTEEV    | 679       |             |
| QY                    | 61               | QAPLAMEAGIQHIELEKSGKGFSLDYOPDIPASVYIIRSLVPGIAEKORLLP       | 120       |             |
| DB                    | 680              | QAPLAMEAGIQHIELEKSGKGFSLDYOPDIPASVYIIRSLVPGIAEKORLLP       | 739       |             |
| QY                    | 121              | GRLMFPVNVNLENSLSEAEALKCAPSGTAVIGVAKPLPSPEEGVYSAKEDSLTYP    | 180       |             |
| DB                    | 740              | GRLMFPVNVNLENSLSEAEALKCAPSGTAVIGVAKPLPSPEEGVYSAKEDSLTYP    | 799       |             |
| QY                    | 181              | HSCEPGLADKPLFRADLALVGTNDALVDETFESPYPENDSIYSTQASISLHGSSC    | 240       |             |
| DB                    | 800              | HSCEPGLADKPLFRADLALVGTNDALVDETFESPYPENDSIYSTQASISLHGSSC    | 859       |             |
| QY                    | 241              | GGGLNGLSSSPSPKDYENSCDPLVLAHMSLEELYTONLEROEDNPSVDISGRPS     | 300       |             |
| DB                    | 860              | GGGLNGLSSSPSPKDYENSCDPLVLAHMSLEELYTONLEROEDNPSVDISGRPS     | 918       |             |
| QY                    | 301              | GFTINDYTPANALIEQYECENTIVMTESHLPSEVISAEPLVLPDSAGKSEHLEQSS   | 360       |             |
| DB                    | 919              | GFTINDYTPANALIEQYECENTIVMTESHLPSEVISAEPLVLPDSAGKSEHLEQSS   | 974       |             |
| QY                    | 361              | LACNAECVLAQNVSKESFERTINIAKGNSSLGMTVYSAKNGKGMVIRSIHGAISRDGR | 420       |             |
| DB                    | 975              | LVSDEASVTLQSKQAFERTVYIAKGNSSLGMTVYSAKNGKGMVIRSIHGAISRDGR   | 1034      |             |

[illegible]

```

RESULT 2
T30259
multiple PDZ domain protein - mouse
C:Species: Mus musculus (house mouse)
C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 21-Jul-2000
C:Accession: T30259
R:Simpson, E.H.; Suffolk, R.; Jackson, I.J.
Genomics 59, 102-104, 1999
A:Title: Identification, sequence, and mapping of mouse multiple PDZ domain protein g
A:Reference number: z20797; MUID:99328525
A:Accession: T30259
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-2055 <Sim>
A:Cross-references: EMBL:AJ131869; NID:94150877; PIDN:CAA10523.1; PID:94150878
A:Experimental source: strain C57/BL6 x CBA F1; whole brain
C:Genetics:
A:Gene: mpdz
A:Map position: 4

```

| Query Match | Similarity | 80.4%   | Score        | 5617:      | DB 2: | Length | 2055: |
|-------------|------------|---|--------------|------------|-------|--------|-------|
| Best Local  | Similarity | 77.1%   | Pred. No. 0: |            |       |        |       |
| Matches     | 1112:      | Conservative  | 114:         | Mismatches | 139:  | Indels | 78:   |
|             |            |   |              |            |       |        | Gaps  |
| QY          | 1          | MYCCRTVPPTTQSELDSDLCDIELTEKPHVDLGEFFISSSETEDEPVLMTADAGOSTEEY  | 60           |            |       |        |       |
| DB          | 621        | MYCCRTVPPIALSEMSDSDINDLELTERPHLDGFISSSETEDEPMLMSVDQAEEI       | 680          |            |       |        |       |
| QY          | 61         | QAPPLAMEAGIOHIELEKSGSKGSLGSLDYODPDPASTVYIIIRSLVPGGIAEKDGRLLP  | 120          |            |       |        |       |
| DB          | 681        | QPLPLAMEAGGOSIELEKSGSRGSGSLDYODPDPANVIVIRSLVPGGIAEKDGRLLP     | 740          |            |       |        |       |
| QY          | 121        | GRLMEVNDVNLNENSLGEAEVALKGAPSGTVIRIGAKPLPLSPREGYSAEDSPLYRP     | 180          |            |       |        |       |
| DB          | 741        | GRLMEVNDIINENSLTEAEVALKGAPGVMRIGAKPLPLSPREGYSAKEDATLGP        | 800          |            |       |        |       |
| QY          | 181        | HSCEBEGADKDLFRADLALVGTNDADLVDESTFESPYSPENDSIYSTQASILSLHGSSC   | 240          |            |       |        |       |
| DB          | 801        | HACKESGJSDKALFRADLALIDTPDESIAESHFESQSPDNDYSTQASIFSLHDGTC      | 860          |            |       |        |       |
| QY          | 241        | GGGLNMGSSLPSSPRDVIENSCDPLVDLHMSLEELYTONLEKDENTPSYDISNGPMS     | 300          |            |       |        |       |
| DB          | 861        | SGGMNMGPSLSSPPRDVYSSS-EVVLGLHLSTLEELYTONLQROHAGSSSDMKPTPF     | 919          |            |       |        |       |
| QY          | 301        | GFTINDYPAANKIEOQYECCENTIVMTESHLPSEVISASLEPLSPSAGKSGEHLLEQSS   | 360          |            |       |        |       |
| DB          | 920        | GFXISDYTTNNGEQKYQANVPANPHSOLPSN-ISTFELASL---AAVARKYLTLDQIF    | 975          |            |       |        |       |
| QY          | 361        | IACNAECVLMQNVSKSEFERTINIAKGNSSLGMTVYSANKDGLMIVRSIIHGAISSRDGR  | 420          |            |       |        |       |
| DB          | 976        | LASDAESVTLQSGMGAFAFERPTYIAKGSSSLGMTVYSANKDGLGIVIRSIIHGAISRDGR | 1035         |            |       |        |       |
| QY          | 421        | IAIGDCILSINEESTISVTNQAQAMLRHSLIGDPDKTTYPAHELEEFKISLGQOSR      | 480          |            |       |        |       |
| DB          | 1036       | IAYGCILSINEESTISLTNQAQAMLRHSHSLIGDPDKITYPAHELEEFKIRFVGQOAG    | 1095         |            |       |        |       |
| QY          | 481        | VVALDIFSSYTRDIPELPEREGEGEESSELONTAYSMNNOPRYVLTMEPKSLGISIV     | 540          |            |       |        |       |
| DB          | 1096       | IVALDIFSSYTRDIPELPEREGEGEESSELONAAISSQPRVELMREPKSLGISIV       | 1155         |            |       |        |       |
| QY          | 541        | GGRGMSRLSNENVMKGFIEKHVLEDSPAGKNGTGLPGDRIYE-----               | 583          |            |       |        |       |
| DB          | 1156       | GGRGMSRLSNENVMKGFIEKHVLEDSPAGKNGTGLPGGRITIEVGDMDLRDASHEDAVYA  | 1215         |            |       |        |       |
| QY          | 584        | -----ADSQSES  | 590          |            |       |        |       |
| DB          | 1216       | IRKAGNPVVFVQSLINRPRKSPLESLPLSKYFSSTNPADSLQJTTQDQASQSES        | 1275         |            |       |        |       |
| QY          | 591        | EPKAPPLGVPPPPPSAFEMSGSDHTQSSAKSISQDVOKDEDEFGYSMKIRIRRYGLTGE   | 650          |            |       |        |       |
| DB          | 1276       | EPKAPPLGVPPSPSASFEMSGSDAQPSATVYSEDEKDEDFGYSMKIRIQRYGLTGO      | 1335         |            |       |        |       |







A:Molecule type: mRNA  
 A:Residues: 1-628 <DHO>  
 A:Cross-references: EMBL:AF034746; NID:g3041880; PIDN:ANC40076.1; PID:g3041881  
 A:Experimental source: brain  
 C:Genetics:  
 A:Gene: lnx

Query Match 8 4%; Score 586; DB 2; Length 628;  
 Best Local Similarity 29.0%; Pred. No. 5,4e-25;  
 Matches 187; Conservative 114; Mismatches 237; Indels 106; Gaps 23;

782 KNYQHLEPKD-----QGLL--GIAISEDTLSGVITKSLTEHGVAATDGRKVGQDILA 834  
 Db 33 KASHHGLTKDKRRKRDQDCPCGCASLMATTL-----PEVSAAT-----ISL 76  
 Qy 835 VDDEIVVGVPIEKISLTKAKTVKRTIHAENPDSQAVP---SAGAAAGEKKNSSQS- 890  
 Db 77 MTEDEPLDNP--AYVSVDEGEPEVANSSDGSRNTRARPERSTWRSRFFKINRALS 134  
 Qy 891 LMAPGSGSEPESEI-----RNTSRSTPAIFASDPATCPIIPGCE--TTEISKRTGLGS 945  
 Db 135 LRRTKSGSVVANHVDGRDNSENTEVPEVF--PRLFHLIPDDELTSIKINRADPSESLS 191  
 Qy 946 I--VGSDFLLGAFIIEHYEEGACCKDGLMAGDQILENGIDLRKATHDEAINVLRQT 1003  
 Db 192 IRLVGSSETPVHIIITQHITRGDVIRADGRRLPGDIIILKVNMDISNPHNVAVRLRQP 251  
 Qy 1004 PORVRLTLRYDE-----APYKEEYVODTLTIELQK-KPGKGLSLSVGRNDTGYE 1053  
 Db 252 CVLRLTLVREQKFRSRNANHVPDSYGRPDSPFVILNKSSPEQGIKLVVRVDEPGVE 311  
 Qy 1054 VSDIVVGGIADPDGRILQGDQILLVNGEDVRNASSQAVAAALCKSLGTYLEGRKAC 1113  
 Db 312 IFNVLVGVADRHGQLEENDRVLAINGHDLRFESPSAAHLQASERRVHLVSR----- 366  
 Qy 1114 FHSERRPS-----QTSQVBSGLSSTFPLSGSSSTLESSSKKNALASEIGLRTVEK 1168  
 Db 367 --QVRSSPDIPQEAQMISNGQS-----PQGR-----NTAKRPAATCHE---KVYSV 411  
 Qy 1169 KKGPTDSLGISIAGVSGPLGDPVIFITAMHPTGVAAGTQKRLVGRIVITIGTSTEGMT 1228  
 Db 412 WKPPSSSLMTVGGASHHEMDPIYIVISVEPGVISRQRIKTGDIILLVNGIELTEVS 471  
 Qy 1229 HNGAVLLKNASSIEMOVAVAGDGVVGHNEBPASSSLSTFGLTSTSFQDDLRP-- 1286  
 Db 472 RTEAVAILKASPSVVKAL--EVK-----EQEAQEDCSPALDS---NNHVTTPGD 518  
 Qy 1287 -----QCKSTLERGPDG-LGFSIVGVGSPHGDLPYIKVTFAGKAAS 1329  
 Db 519 WPSWVWMLLEPOLYLCNKADVILRRNTAGSLGCIYGGIEIYSGNRPFFIKSIVEGTPAT 578  
 Qy 1330 EDGRKRGDQIIIVANGOSLEGVTHEEVAAILKRTKGTVTLMVL 1373  
 Db 579 NDRIRGCDILLAVNGRSTSGMIACIARMLKELKGRITLTIAS 622

RESULT 6  
 T23160  
 hypothetical protein K01A6.1 - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C:Date: 15-Oct-1999 #sequence\_revision 15 -Oct-1999 #text\_change 18-Aug-2000  
 C:Accession: T23160  
 R:Cottage, A.  
 submitted to the EMBL Data Library, January 1996  
 A:Reference number: Z19701  
 A:Accession: T23160  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-1012 <MLT>  
 A:Cross-references: EMBL:268750; PIDN:CAA92965.1; GSPDB:GN00022; CESP:K01A6.1  
 A:Experimental source: clone K01A6  
 C:Genetics:

A:Gene: CESP:K01A6.1  
 A:Map position: 4  
 A:Intron: 107/1; 151/2; 183/2; 270/2; 316/3; 410/3; 628/3; 922/3; 971/3.  
 C:Superfamily: WW repeat homology  
 F:131-168/Domain: WW repeat homology <WWR>

Query Match 7 6%; Score 534; DB 2; Length 1012;  
 Best Local Similarity 24.1%; Pred. No. 9,6e-22;  
 Matches 208; Conservative 123; Mismatches 298; Indels 234; Gaps 31;

572 NGLPKGRDIVEAPSSSESEPEKAPLCVY-----PPPPAFAMGSDHTQSSAKIS 624  
 Db 317 SGTLK-----SSSPRDSGDSPTTRKRGDPERATSSADVDHSHKMSRSS 365  
 Qy 625 QDVDEDEFGYSKNIERYGTLTGELHMELEKSHGSLGLSLAGNKRSMRSVFI--VG 682  
 Db 366 NPLFTTDP-----ARLGELISTKIKYKAGKAGFTLIGNDSSRDEFTQVKS 413  
 Qy 683 IDPNGAAGDKRLQADELLEINGQILYGRSHQNASIIRKAPSKYKIFIRNKDAVNO- 741  
 Db 414 VLSGPGAAANGVLRSGDILLVRNGLLLGATQKEACDVEVAIP-----VNDA 460  
 Qy 742 --MAVCPGNAVEPLPSENENLQNKETEPTVTTSDAAVDSFRKNVOHLEPKDQGLGIA 799  
 Db 461 VDIQVGRGYELFDIPAN-----RIVTENYAAAKS-RDLHEIDIFGSEGFET 508  
 Qy 800 ISEEDTLGVIIKSLTEHGVAAATDGRKLVGDQILAVDEIVVGVPIEKISLTKAKTV 859  
 Db 509 IA--DNLNGORIKIL---FPSQCPMLMEGDTIVELDGNVRIPIHTQLVMDLREPIGY 563  
 Qy 860 --KLTHAENPDSQAVPSAAGAAAGEKKNSSOSLMP-----QSSSPESIR 905  
 Db 564 RGLVVKRSGSPKTRSTPFAAFRYGEPQTNMDSAPLVRSKTPAEROTSREEDQNYR 623  
 Qy 906 NTSRSTPAIFASDPATCPIIPGCEETTEISKRTGLGISVGSDFLLGAFIIEVYEE 965  
 Db 624 NT-----LQRPVAVTSEWEMSSAIPASRRP-----STTLTG- 657  
 Qy 966 GAACKDRLMAGDQILEVNGIDLRKATHDEAINVLRQTPORVRLTYRDEAPYKEEVC 1025  
 Db 658 -----ATPNYIPLSQ-----NQKPSD 674  
 Qy 1026 TLTIELQKPKGGLSIY--GRNDTGVFVSDIVKGIADPDGRILQGDQILLVNGEDVR 1084  
 Db 675 LITVSLIRKP-VGEGFRLGVEESKTPLSVGQIVIGAAEDGRLOEGDEIVEIDSHNE 733  
 Qy 1085 NASQEAVALLKSLCT--VTLEVGRIKAGPFHSERRPSQTSQVSGSLSTFPLSGSS 1142  
 Db 734 GASHSEAVVLLERAAKNHVKLY-----RRPSRTDPARGSLNS-----AGPS 777  
 Qy 1143 TSELESSSKKNALASEIOGLRTVEKKKGPPTDSLISINGV---GSLGDPVIFITAMH 1199  
 Db 778 -----GSYDVLLHRNENDGFGVLLMSQHKNSSTVGOI-----Q 811  
 Qy 1200 PPGVAAOTKIRVGRDRIVITIGTSTEGMTHTQAVNLLKNASSIEMQY-----VAGDVS 1255  
 Db 812 PGSPARCGRLSGVDRVIANVIGDILSLSHPTISLTKDSGLSVRLTIPMTAGVLP 871  
 Qy 1256 VT-----GHHO-----EPASSLSFTGLTSTSFQDDLRP--PPOCK 1289  
 Db 872 VSAITLNRFTMGHYESNNGLPPEPPSYVEKHPPEYLAIFDGLS-----INDRSMNGNLI 927  
 Qy 1290 SITLERGPDGLGSIYGVGSPHGDLPYIKVTFAGKAASEGRKRGDQIIIVANGOSLE 1349  
 Db 928 DVTLEGTGKGFGFSING--GOEFGSMPLFVLRITADGPAKADRLQVGDQLTTINGQSTK 985  
 Qy 1350 GVTHEEVAAILKRTKGTVTLMVL 1372  
 Db 986 GMSHDAIRIIRKO-HTMVNLITVL 1007

RESULT 7



```

Db 326 PEAVYRTSTTRKKARISDGSIALDIFGPOKMDPIYTRRELPTS---SALSSALDRR 382
Oy 268 DLHMSLEELYTONLERODENTPVSIDSMGASGFTINDYTPANAIEQOYEGENTIWTE 327
Db 383 EKQKTL-----QVLRAMNVEEP-----VRKTYHGDVFSRS 415
Oy 328 SHLPSEVISAEPLPSLPASGK-----GSEHLEKQSSL-----ACN 364
Db 416 SESPSTISSESDFOYRREASKRFESSSGLPVDETLGOGOSQPSROYETPFEGNLI 475
Oy 365 AECVMLOANSKESFERTINIANKNSL-----GMTVSN-----KQDLGMIV-----406
Db 476 QE-IMLKROBELMOLQAKMALRQSLSLPGDTIKASMLDITRDLREIALETANTQK 534
Oy 407 -----RSTIHGSAISRDRGRIAGDCILSINEESTISVTNAO- 442
Db 535 LRFNFGPEPVKMTIEPFIIDLPRSLITFKKKNEDNRKRYN--IMLLNGORLELTCDTK 592
Oy 443 -----ARAMLRRHSL-----IGPDITITYPAHLEB-----469
Db 593 ICKDVEPMVAVHIGLVEHHLFALATLKDNEYEFVDPDLKLTVPARGMKEEPKKTKATV 652
Oy 470 -----FKI-----SLGQOS-----478
Db 653 NPTLEFRIFEMDVSLLQHTLTLCHQYIQLKRDILEERNHCDDETLASLALQAEYG 712
Oy 479 -----GRVA-ALDIFSSYTGRIPELPEREAGE-ESLQNTAVS 517
Db 713 DYQPEVHGVSYFMEHLYLPAVMEKLDL--SYIKELPKLHNTYVASEKETELFEFLKYC 770
Oy 518 NMNQPRVELMR-EPK-----SLGISIV-----GGRGMSRLSGEVMRGIIFK 561
Db 771 ORLTERGVHHRVHPKKSOTGILLGVCSKGVLFVFNHNGVRLVLRFPRREKKKSFSF 830
Oy 562 H--VLEDSFAG--KNGTLKGRDRIVEAPSOSESEPERAPLCSVPPPSAFEMGSDHTQS 618
Db 831 KKTLLQNTSGIKHGFOTQMSKICQYLLHCSYQHKFOL-----QM 871
Oy 619 SASKIQDVKDEDEFGYSKNTIR--ERYGTLGELMHMIELEK-GHSGGLSLAGNDRSR 675
Db 872 RANQSNQDODIDVLHRRKMSIYSPER-----EITLVNKKAKKGLGOIIGGKMG 925
Oy 676 --MSVTVIGDIPGAGKQDRLOLADLELEINCOILYGRSHQVASSIKCAPSKVIIF 733
Db 926 LDGRTISSVAPGPRADLDGCLKPRDRLISVNSVLEGVSHHAIETLQNAPEDEVLVIS 965
Oy 734 RNKDAVNOAIVCPGNAVEPLPSNSENLOKTEPTVTTSDAVALDSFKN-----VOHL 787
Db 966 QPEKISKVSPSTVH-----LTNEMKNVKKSS-----YMQDSALDSSSKDHMSRGLRHI 1037
Oy 788 ELPK--DQGL--GIAISEEDTLSSVYIISLTHGVAALDGRKAVDQDILAUNDEIVG 843
Db 1038 SENSFGPSGLRSGSLSDQSRTESASLSQSQNGFFAS-----HLDQDTV--QEOHGS 1090
Oy 844 PIKFIISLKTAKMTVKLITHAENPDSQAVPSAAGASEKK-----NSOSLW 893
Db 1091 PPSVSVS-----KATEKETTDNSQSKTKKPGISDVTVSDRSDSDMEATVYSSSDHOT 1145
Oy 894 POGSPEPESIRNTSSSPALFASDPATCPRIIGCEETIETISKGTGLSLIVGSDTL 953
Db 1146 PKQ-----ESSSVNTSNKMNKFTSSSPK-----PGDIFVELAKNNSIGISVTGCVNTS 1198
Oy 954 L--GATIIHEVEGAACKDGRIMAGDQILEVNGIDIRKATHDEAINVLKQTPQRYRL 1011
Db 1199 VRHGCIYVAVIPOGAESDGRITHKDRVLAVNGVSLGATHKQAVETLNTQOVYHLL 1258
Oy 1012 YRDEAVYKEEV-----C-----DRLTELEKRP 1036
Db 1259 EKQOSTSKEHVYTPQCTISDONAGQGEKAKTTQVKDYFVLEENTFEVKLRN-S 1317
Oy 1037 KGLGLS-----IVGRNDTVGVSDIYVKGIALDPDGLILOGDQILLVNGEDVRNASOE 1089

```

```

Db 1318 SGLGFSFREDNLEIPEQINASTIVRYKKLEPGQPAAESKIDVDYILKNGASLKLGSQ 1377
Oy 1090 AVAALLKCSLGTVTLVEGRKAG-----PFRSE-----1117
Db 1378 EVISALRGTAPEVFLLCRPPGVLEIDTALLTFLQSPAVQLPNSSKDSQPSCEQST 1437
Oy 1118 -----RRPQTSQVSEGS-----LSSFTF-----1136
Db 1438 SSDNEMSDSKKQCKSPSRSDSYSDSGSGEDDLVTAPANISNSTWSALHQTLSNWS 1497
Oy 1137 -----PLSGST-----SESLESS 1150
Db 1498 QASHHNAKSDIEDTICTMFYYPQKIPNKPEREDSNPPLPDMAFGOSYOPSSASS 1557
Oy 1151 SK-----KNALASEIGLR-----TVEKKGPDSLCISLAG-- 1183
Db 1558 SMDKYHHIHSPTROENTPLKNDLENHLEPELEVELLITLILIKSEKSLGFTYKQNO 1617
Oy 1184 -VGSPLGDYPIFLAMHPRTGVAAGQOKLAVGDRIVTIGSTEGMTHQAVMLKNASGS 1242
Db 1618 RIGCYVHDV-----IQDP--AKSDGRLPGRDLKLVNDVTNMTHTDAVNLRAASKT 1669
Oy 1243 IEMQVYAGDVAVYTGHNQE-PASSLSFTGLTSTIFQDDLGPPQCKSITEERGPDGL 1301
Db 1670 VRL-----VIGRVLELRIPML-----PHLLPDTITLCKNEELG 1703
Oy 1302 FSTVGGSPPHGDLPIYKTVFAKGAASEDRKRGDQIIAVNGSLSCVTHEEA-----1356
Db 1704 FSLCGGHSDSLY--QVYIYSDINPRVAALTEGNLQDLIDVHYVNGVSTQGMTEVEVNRALD 1761
Oy 1357 ---VALIKRTKGTVTLM 1370
Db 1762 MSPLSLVLKATRNLDLPV 1779

```

```

RESULT 9
167629
protein tyrosine phosphatase (PTP-BAS, type 2) - human
C:Species: Homo sapiens (man)
C:Date: 29-May-1998 #sequence,revision 29-May-1998 #text_change 16-Jul-1999
C:Accession: I67629
R:Maekawa, K.; Imagawa, N.; Nagamatsu, M.; Harada, S.
FEBS Lett. 337, 200-206, 1994
A>Title: Molecular cloning of a novel protein-tyrosine phosphatase containing a membr
A:Reference number: I53483; MUID:94116679
A:Accession: I67629
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-2466 <RES>
A:Cross-references: GB:D21210; NID:9452191; PIDN:BA04751.1; PID:9452192
F:574-866/Domain: protein-tyrosine-phosphatase, nonreceptor type 13; GLGF domain homolog
F:1354-1430/Domain: GLGF domain homology <GLG2>
F:2218-2437/Domain: protein-tyrosine-phosphatase homology <PTP>

```

```

Query Match 7.3%; Score 509.5; DB 2: Length 2466;
Best Local Similarity 22.3%; Pred. No. 9,8e-20;
Matches 256; Conservative 146; Mismatches 395; Indels 351; Gaps 41;

Oy 510 ELQNTAVSMNQPRVRLWRREPSSKLSGISVGRGMSRLSNG-----552
Db 868 QLOMRARQSNQADIERASFRLNQAESYRGFNNGRAISTGSLASLTNLKLAIVPLSV 927
Oy 553 --EVMNGIFIKHY-----LEDSPAGKNGTL---KPD--RIVEAPSOSESEPERAPLCS 599
Db 928 QAEILKRLCSSELSTLOPLQNSKKEKNDKASWEKPREMSKSYHDSQSLVPHRRNVIV 987
Oy 600 VPPPPSAFAEM-GSDHTQSSAS-----KTSQDVDRKD 631
Db 988 NMEPPQVAVELVKGKPSHMSRSDAESLAGVYTKLNSKSVASVLSNRPERRKHESDSSSIE 1047
Oy 632 EFG--YSWKNIIRERYTLNG---ELHMIIELEK-GHSGGLSLAGNDRSR--MSVTVIGI 683

```

```

Db 1048 DPGQAVYVDLVLRKMSIYSSPEREITLVNKKDAKYGLGFOITIGCEKMGRLDGIFFISSV 1107
OY 684 DPNGAAGDGRLOIDELLEITLNGQLYGRSHONASSIIKCAAPSKYIIIFIRNKDAVNOMA 743
Db 1108 APGGPADLDGCLCPGDRILISVNSVLESEVSHAAIEITLQNAPEVDVTLVISQPEKIKYVP 1167
OY 744 VCPGNAVEPLPSNSENLOKETEPTVTTSDAVAVDSLFKN-----VOHLELPK--DOGG 795
Db 1168 SPPVH-----LTNEMKNYKKS-----YMODSALDSSSDHMHMSRGLNHSNNSGPPSGG 1219
OY 796 L--GAISEEDTLGCVITIKSLTEHCVATDGRKLYGDOILAVDEIVVGYPIEKFIISLK 853
Db 1220 LREGSLSSQDSRTESASLSQSQVNVGFAS-----HLGDQW---QESQHGSPSPVYS--- 1269
OY 854 TAKMVKTLIHAENPDQAVPSAAGASGEKK-----NSSQSLMPPSSGPES 903
Db 1270 --KATEKETFDSDNSQSKKKKKEISDVITYSDRGSDMDENATYSSQDHQTPKQ--ESSS 1324
OY 904 IRNTRSSSTPAIFASDPATCPITPCETITELSKRGTLGSLIVGSPDLL--GAFIIE 961
Db 1325 SVNTSNKMNKFTFESSSPK-----PDIFIEVELAKNDLSIGSVGVNTSVRHGGIYKA 1380
OY 962 VYEEBAACKDGRMLAGDOILEVNGIDELKATHDEAINLRQTPQVRYRLTYRDEAPYKEE 1021
Db 1381 VIPQGAESDGRIRHKGDRVLAVNGVLEGATHKQAVETLRMTGQVYVHLLLEKGSPTSKE 1440
OY 1022 EV-----C-----DTLTIEOKKPKGGLIS----- 1042
Db 1441 HVPYTPQCTLSDQNAQGGPEKVKTTQYKQSYFTEENTBEVKLFKN--SSGLGSESRRE 1499
OY 1043 ---IVKRNQDGVFSDIVYKGIADPRDLRGDQILLVNGEDVYNASQEAVALLKCSL 1099
Db 1500 DNLPEQINASTIVRYKKEFGQPAAESGKIDYGVITLVKNGASLKGSLQGEVETSLRGT 1559
OY 1100 GTVILEVRIRAG-----PFHSE----- 1117
Db 1550 PEVELLILCRPPGVLPETIDTALTPLQSPAQVLPNSSKDSOPSCVEOSTSDENEMSDK 1619
OY 1118 ---RRPSOTSQVSEGS-----LSSFTF----- 1138
Db 1620 SKKOCKSSRSDYSIDSSGSEGDLYTAPANISNTWSALHQLLSMNVSQAOQHHEAPK 1679
OY 1139 SGSSST-----SESLSSSK----- 1152
Db 1680 SQEDICTMYYPOKIPKPFEDSNPSLPDPMAPGQSYQPOSASASSMDKYHINI 1739
OY 1153 ---KNALASEIOGLR-----IYEMKGPIDSIGISAG--VGSPLGVP 1192
Db 1740 SEPTROENWPLKNDLENHLEDFELELLITLLIKSEKSLGFTVYTKGNRIGCYVHGV- 1798
OY 1193 IFIAMHPTGVAATOKLRYVDRIYITIGTSTEGTHTQAVNLKLNASISTEMOVVAGCD 1252
Db 1799 ----IQDP---AKSDGRKLKPDRLIKVMDTVDVNTHTHDANLRAASKYVRL----- 1844
OY 1253 VSVVTGHHQE--PASSSLSTGLTSTSTIFQDGLPPQCKSTILERKPDGLGFSIVGYGSP 1311
Db 1845 ---VIGRVLLEIPRIPML-----PHLLPDIITLTCNKKEILGFSLGGHSL 1885
OY 1312 HGDLPIYKTVFANGAASEDRGLKRGDQIIIVNGOSLEGVTHEEA-----VALIKR 1352
Db 1886 Y--QVYIYSDINPSVAIBENLOLDVITHYVNGVSTOGMTLEEVNRALDMSLSVLKA 1943
OY 1363 TKGTVTLLM 1370
Db 1944 TRNDLPVY 1951

```

RESULT 10  
 A54971  
 protein-tyrosine-phosphatase (EC 3.1.3.48), nonreceptor type 13 - human  
 N:Alternate names: Fas-associated phosphatase FAP-1; protein-tyrosine-phosphatase hPTP1F  
 C:Species: Homo sapiens (man)

C:Date: 11-Nov-1994 #sequence,revision 08-Feb-1996 #text,change 01-Dec-2000  
 C:Accession: A54971; A55114; I55955; I53483; S46955  
 R:Barville, D.; Ahmad, S.; Stocco, R.; Shen, S.H.  
 J. Biol. Chem. 269, 22320-22327, 1994  
 A:Title: A novel protein-tyrosine phosphatase with homology to both the cytoskeletal  
 A:Reference number: A54971; MUID:94350988  
 A:Accession: A54971  
 A:Molecule type: mRNA  
 A:Residues: 1-2490 <RAN>  
 A:Cross-references: GB:U21278  
 A>Note: sequence shown follows authors' translation at positions 62-63  
 R:Satas, J.; Claesson-Welsh, L.; Heldin, C.H.; Gopez, L.J.  
 J. Biol. Chem. 269, 24082-24089, 1994  
 A:Title: Cloning and characterization of PTP1, a protein tyrosine phosphatase with s  
 A:Reference number: A55114; MUID:95014139  
 A:Accession: A55114  
 A:Molecule type: mRNA  
 A:Residues: 1-61, 'GS', 64-839, 'D', 841-1055, 1075-1133, 'FH', 1136-1210, 'T', 1212-1383, 1389  
 A:Cross-references: GB:X80289; NID:9515030; PIDN:CMA56563.1; PID:9515031  
 R:Sato, T.; Irie, S.; Kitade, S.; Reed, J.C.  
 Science 268, 411-415, 1995  
 A:Title: FAP-1: a protein tyrosine phosphatase that associates with Fas.  
 A:Reference number: I59595; MUID:95232528  
 A:Accession: I59595  
 A:Status: translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1279-1888 <RES>  
 A:Cross-references: GB:L34563; NID:g806291; PIDN:AAC41755.1; PID:g806292  
 R:Maekawa, K.; Imagawa, N.; Negamatsu, M.; Harada, S.  
 FEBS Lett. 337, 200-206, 1994  
 A:Title: Molecular cloning of a novel protein-tyrosine phosphatase containing a membr  
 A:Reference number: I53483; MUID:94116679  
 A:Accession: I53483  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-61, 'GS', 64-839, 'D', 841-1210, 'T', 1212-1383, 1389-2299, 'QW', 2302-2490 <RE2  
 A:Cross-references: GB:D21209; NID:9452189; PIDN:BA04750.1; PID:9452190  
 C:Genetics:  
 A:Gene: GDB:PTPN13  
 A:Cross-references: GDB:306348; OMIM:600267  
 A:Map position: 4q21.3-4q21.3  
 C:Superfamily: protein-tyrosine-phosphatase, nonreceptor type 13; GLGF domain homolog  
 C:Keywords: alternative splicing; phosphoprotein; phosphoric monoester hydrolase; tyr  
 F:574-868/Domain: protein 4.1 membrane-binding domain homolog <B41>  
 F:1099-1175/Domain: GLGF domain homology <GLG1>  
 F:1373-1454/Domain: GLGF domain homology <GLG2>  
 F:1511-1590/Domain: GLGF domain homology <GLG3>  
 F:1799-1870/Domain: GLGF domain homology <GLG4>  
 F:1893-1967/Domain: GLGF domain homology <GLG5>  
 F:2242-2461/Domain: protein-tyrosine-phosphatase homology <PTP1>  
 F:2413/Active site: Cys (phosphocysteine intermediate) #status Predicted  
 F:2419/Binding site: substrate phosphate (Arg) #status predicted

Query Match 7.18: Score 494.5; DB 1; Length 2490;  
 Best Local Similarity 22.66; Pred. No. 6.8e-19;  
 Matches 243; Conservative 140; Mismatches 371; Indels 319; Gaps 38;

```

OY 530 EPPKSLGISTYGRGMSRLSNGEVMRGIFIKYHLEDSPAKNGTKLPGRIVAPQSGE 589
Db 990 EPPQTYAELVGKPSHOMNSDAESLAGV---TKLNSKSVASLNRKPERKHESDSSSI 1046
OY 590 SEPEKAPLCVPPPPSAFAEMGSDHTQSSASKISQDVNDEDEDEFGYSWKNIRERYGTLTG 649
Db 1047 EDPQAVYLQMT-----MHSSGNSSSQVPLKENDV-----LHKRWSTVSS 1086
OY 650 ---ELHMIIELEK-GHSGLGSLAGNKRDR--MSVFTVIGDIPNGACGDRLOIADLELE 703
Db 1087 PEREITLVNLLKAKYKGLIGFOITIGCEKMGRLDGIFFISSVAPGPAIDLDCCLCPGRLIS 1146
OY 704 INQILYGRSHONASSIIKCAAPSKYIIIFIRNKDAVNOMAVCGNANPELPSENENLQNK 763
Db 1147 VNSVLEGVSHHAAIEITLQNAPEVDVTLVISQPEKIKYVPSTPVH-----LTNEMKNYMK 1202

```



|                       |                  |                    |             |             |
|-----------------------|------------------|--------------------|-------------|-------------|
| Query Match           | 5.88:            | Score 407.5:       | DB 2:       | Length 852: |
| Best Local Similarity | 28.48:           | Pred. No. 8.2e-15: |             |             |
| Matches 155:          | Conservative 68: | Mismatches 212:    | Indels 111: | Gaps 22:    |

[illegible]

C:Species: Rattus norvegicus (Norway rat)  
 C:Accession: T14152  
 C:Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 18-Aug-2000  
 R:Hitao, K.; Hata, Y.; Ide, N.; Takeuchi, M.; Irie, M.; Yano, I.; Deguchi, M.; Toyoda, A.  
 J. Biol. Chem. 273, 21105-21110, 1998  
 A:Title: A novel multiple ptd domain-containing molecule interacting with N-methyl-D-asp  
 A:Reference number: Z17889; MUID:98361985  
 A:Accession: T14152  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-1277 <H1R>  
 A:Cross-references: EMBL:AF034863; NID:g3411052; PID:g3411053; PIDN:AAC31124.1  
 C:Genetics:  
 A:Gene: S-SCAM  
 C:Function:  
 A:Description: may assemble receptors and cell adhesion proteins at synaptic junctions  
 C:Superfamily: WW repeat homology  
 F:302-339/Domain: WW repeat homology <WM1>  
 F:348-385/Domain: WW repeat homology <WM2>

Query Match 5.7%; Score 398.5; DB 2; Length 1277;  
 Best Local Similarity 19.9%; Pred. No. 5.1e-14;  
 Matches 221; Conservative 146; Mismatches 333; Indels 413; Gaps 46;

```

QY 30 PHVDGEFISSEDEDPVLAAMDAGOSTEEVOAP-----LAMMEAGIHIETLEK 78
DB 393 PTELEG-----TKPLQAPGFRKPLRTTRDASQKGFLLSTTLKK 431

QY 79 GSKGLGFSILDYODPIDPASTVILIRSLVPGGIAEKGRLLPQDRLMFVNDYLENSSLE 138
DB 432 SMNGGFTIIGDEP---DEFLQYKSVIPDGPAAQDKMERTGDIYVINEYCVLGHTHA 487

QY 139 EAVEALKAPSG-TYRIGYAK--PLPLSPEGVSAKEDSFYIPHSCEAGLADKPLER 195
DB 488 DYYKLEQSPVIGOSVNLVYRCGYPLPFEDEPDANS-----WVPP-----LAIMERP--- 533

QY 196 ADLALVGTADALVDESPESPSPENDSIYQASILSLHSSCGDGLNTYSSLPSP 255
DB 534 -----PPVMMNGRINNEYLEYEYISRTSOSVDP-----ITDRPP 566

QY 256 KDVIENSCDPVLDLHMSLEELYTONLIERODENTPSVDISMGPASGFTINDYTPANAIDQ 315
DB 567 H-----SLH-----SM-PADQQLDGTTPP----- 585

QY 316 QYECENTIVWTESHLPSEVISSALPVLPSDAGKSGEHLDDQSSSLACAEVMAQNSK 375
DB 586 -----VHDDNVSYA-----SSGATQAEILM----- 604

QY 376 ESFERTINAKNSSLGMYVANSKDGIMIVRSIINGAISNDGRITADCTLSTINEST 435
DB 605 -----TLTIVYKAKGFGFTIADSP--TGQVAKOILD--IQCGPGICEBDLVEIQNV 654

QY 436 ISVTNAQARMLRRHSLIPDIKITVYPAHLKEFKISLGQSGRVMALDIFSSYGRDI 495
DB 655 QNLSTHE-----VVDIL-----KCC 669

QY 496 PELPEREGEGESEL--ONTAVSNMNPOR-RVELMR--EPSSKLSISVGGRMGSR 548
DB 670 P-----VGSETSLIIRGGEFFSPMKTPKPMVDNEMENQSPQSLSPAPV----- 713

QY 549 LSNGEVNRGIFIKHVLDESPACKNGTLKPGDRIVEAPSSSES-EPKALCLSVPPRPPSA 607
DB 714 -----PQSLPPPALHRSF-----PDSTEFADPK-----PDPE 744

QY 608 PAEMGSDHTOSSASKISQDVDEKDEFGYSKMNIRERYGTLTELMHTELEKSGGLSL 667
DB 745 LYEKSAIYESNOQVPRRSFRDSSGPRYKEL-----DVHLRRME--SGGFRPI 792

QY 668 AGNKRDSRNSVPIYGDIPGAAGKDGRLQIADELLEINGQILYGRSHQANASIIKCAPSK 727
DB 793 LGG-DEPGQPLILGAVIANGSADRDGRLHPGDELYVDGIPVAGKTHRYVIDLHMHAARN 851

```

QY 728 VKI-IFIRNKQAVNOMAVCPGNAVEPLPSNSNLONKETEPIYVTSDAVDSLSPKNQV 786  
 DB 852 GGVNLTFRKK-----VLGG-----EPCPNGRSPGSVSTHSSPSRSDYA---TYANSNH 898  
 QY 787 LELPKDQG-----LGIAISEEDTLGVIKISLT--EHGVAAT----- 822  
 DB 899 -AAPSNMNSPREGFASHSLQSDYIIHKKENSGFVYISSLNRESGATTVIPKIGRI 957  
 QY 823 -DG-----RLKVDQILAVDEIVGPIEKETSLKTAKMTVKLT---HAENPDS 870  
 DB 958 IDGSPADRCALKKVDRIILVNGOSIIMPHADIVKLKDAOLSTLRIIPQEEINPTFS 1017  
 QY 871 QAVPSAGASGEKKN-----SSQSLMVPQSSPEDESIR---N 906  
 DB 1018 -----APSEKQSPMAQHSPLAQHSPLAQSPPAIPNSPVAQAPAPQPLQLQGHEN 1069  
 QY 907 TSSRSTPA-----IFASDPATCPPIPGCE----- 930  
 DB 1070 SYRSEVKANQDVKPDIRQPFPTDYRQRPPLDYRQPGGDVSGSPPLDYRQHSPTDQYPLS 1129  
 QY 931 -----TTIEISKRTGIGLSTIVGSDTLGAFIIEHYEERGACKDGRLNAGDI 980  
 DB 1130 DYROPQDEDFYFVDMKGAKEGFGFSIRGREGYKMDLYVL-RLAEDGPALIRNGRMVGDQI 1188  
 QY 981 LEVNGIDLRKATHDEAINVLRQTPQVRRLTYR 1013  
 DB 1189 IEINGESTDMTHARAILIKSGGRVRLLLKR 1221

RESULT 15  
 G01974  
 C:Species: Homo sapiens (man)  
 C:Date: 21-Dec-1996 #sequence\_revision 06-Jun-1997 #text\_change 21-Jan-2000  
 A:Accession: G01974  
 R:Kim, E.; Cho, K.  
 submitted to the EMBL Data Library, July 1995  
 A:Reference number: G08966  
 A:Accession: G01974  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-870 <KIM>  
 A:Cross-references: EMBL:U32376; NID:g1463025; PID:g106790  
 C:Genetics:  
 A:Gene: chapsyn-110  
 C:Superfamily: discs-large tumor suppressor; GLGF domain homology; guanylate kinase h  
 F:198-276/Domain: GLGF domain homology <GLG2>  
 F:543-601/Domain: SH3 homology <SH3>  
 F:681-858/Domain: guanylate kinase homology <GKI>

Query Match 5.7%; Score 397.5; DB 2; Length 870;  
 Best Local Similarity 28.1%; Pred. No. 3.1e-14;  
 Matches 154; Conservative 69; Mismatches 205; Indels 121; Gaps 23;

```

QY 868 PQSQAVPSAAGASC-----EKKNS-----QSLMVPQSSPEDESIRNTRSS 911
DB 24 PDDHSLPRLTHEVRGPELVHSEKMLQIENYGVYLOSHISPLASAPV-IIVMTDL 82
QY 912 TPAIFASDPATCPPIPGCEFTIE-----ISKRTGIGLSTIVGSDPT-LLG---AFIIEHY 962
DB 83 -----TIPYVNGTIEIEFEETILERKNSGLGFSIAGCTDNPHIGDDPGFITIT 132
QY 963 YEEGAACKDRLMAGDQILLEVNGIDLRKATHDEAINVLRQTPQVRRLTYRDEAPYKEE 1022
DB 133 IKGAAAEEDGRILRVNDICIRVNEVDVSEVSHSKAVALKEASIRALTY-RRRRPILFTV 191
QY 1023 VCDITLIELOKPKGKGLSTYK-----RNDTGVPVSDIYKGGIADPDGRLIGDQIL 1076
DB 192 V---EIKLFKRP-KGLGFSIAGVGNGNHIIPDINSIYVKIIDGGAAGKDGRLQYGDRL 246
QY 1077 LVNGEDVRNASQAVALLKCSIGVTLTEVGRIKA-----GP---PHSERRPQSOTSOY 1126

```



```

Db 247 MWNNSLEEVTHEEVAAILKNTSEVYVYLVKGNPTTIYWTDEYCPDITHSYSPMENHL 306
QY 1127 SEGSLSPFTPLSGSSSTSESLSSKKNALASEIQULRTVEMKKGPPTSLGISTAGVGS 1186
Db 307 S-GNNGTLEIXTSLPPISPGRYSPIPKMLVDD-DYTRPE---PYSTYVKKLQDKPAS 360
QY 1187 PLGDVPI-----FTAMHPTGVAATOCLRVGDRIYTCGTSTEGMTHTQAVNLLKNASG 1241
Db 361 PRHYSPECDKSF-----LSAPYSHYHLG-----LPPDSEWTSHQSHSTATROPSM 407
QY 1242 SIEMQVYAGDVSVYTGHHQEPASSLSFTGLTSTSIFQDDLGPPQCKSTILERGPDGLG 1301
Db 408 TLQRAVSLG-----EP-----RKVYLHKGSTGLG 432
QY 1302 FSIYGYGSPHGDLPYVYKTVFAKKAASEDGRKKRQDIIAVNGOSLEGVTHEEVAAILK 1361
Db 433 FNIVG---EDGE-GIFYSFILAGPADLSELQRGDQIILSVNGIDLKASHEDQAAAAALK 488
QY 1362 RTKGTVTLM 1370
Db 489 GAGQTVTII 497

```

Search completed: July 12, 2001, 14:41:51  
 Job time: 199 sec

7

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2001, 14:39:47 ; Search time 21.71 Seconds

(without alignments)  
2166.411 Million cell updates/sec

Title: US-09-502-698-1

Perfect score: 6990

Sequence: 1 MWCRRTPPTTQSELDSD.....EEAVAIKKRTKGTITLVLS 1373

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 3425486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_39:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query Match | Length | ID           | Description        |
|------------|-------|-------------|--------|--------------|--------------------|
| 1          | 509   | 7.3         | 2485   | 1 PTND_HUMAN | Q12923 homo sapien |
| 2          | 405.5 | 5.8         | 852    | 1 DTG2_RAT   | Q63622 rattus norv |
| 3          | 399   | 5.7         | 817    | 1 DTG3_HUMAN | Q92796 homo sapien |
| 4          | 397.5 | 5.7         | 870    | 1 DTG2_HUMAN | Q15700 homo sapien |
| 5          | 393   | 5.6         | 849    | 1 DTG3_MOUSE | P70175 mus musculu |
| 6          | 391   | 5.6         | 849    | 1 DTG3_RAT   | Q62936 rattus norv |
| 7          | 384   | 5.5         | 767    | 1 DTG4_HUMAN | P78352 homo sapien |
| 8          | 383   | 5.5         | 724    | 1 DTG4_RAT   | Q62108 mus musculu |
| 9          | 382   | 5.5         | 960    | 1 DTG1_MOUSE | P31007 drosophila  |
| 10         | 379   | 5.4         | 911    | 1 DTG1_RAT   | Q62696 rattus norv |
| 11         | 378.5 | 5.4         | 904    | 1 DTG1_HUMAN | Q12959 homo sapien |
| 12         | 377   | 5.4         | 904    | 1 DTG1_HUMAN | Q14005 homo sapien |
| 13         | 283   | 4.0         | 631    | 1 I116_HUMAN | Q07157 homo sapien |
| 14         | 243   | 3.5         | 1736   | 1 Z03_HUMAN  | Q95049 homo sapien |
| 15         | 242   | 3.5         | 933    | 1 Z03_HUMAN  | Q62683 canis famli |
| 16         | 240   | 3.4         | 898    | 1 Z03_CANFA  | P33447 mus musculu |
| 17         | 237.5 | 3.4         | 1745   | 1 Z01_MOUSE  | Q94941 mus musculu |
| 18         | 235.5 | 3.4         | 905    | 1 Z01_MOUSE  | P55196 homo sapien |
| 19         | 231   | 3.0         | 1816   | 1 AFE_HUMAN  | Q09053 gallus gall |
| 20         | 205.5 | 2.9         | 839    | 1 AFE_RAT    | Q09053 gallus gall |
| 21         | 197.5 | 2.8         | 3562   | 1 PCGV_CHICK | Q12888 homo sapien |
| 22         | 185.5 | 2.7         | 2109   | 1 PCGA_CHICK | Q12888 homo sapien |
| 23         | 184.5 | 2.6         | 1972   | 1 P531_HUMAN | Q12888 homo sapien |
| 24         | 182.5 | 2.6         | 982    | 1 ABPA_CAEEL | Q12888 homo sapien |
| 25         | 182   | 2.6         | 206    | 1 OM25_RAT   | Q94941 mus musculu |
| 26         | 178   | 2.5         | 2124   | 1 PCGA_RAT   | Q94941 mus musculu |
| 27         | 177.5 | 2.5         | 571    | 1 AFB3_MOUSE | Q08888 mus musculu |
| 28         | 177   | 2.5         | 837    | 1 AFB3_HUMAN | Q08888 mus musculu |
| 29         | 176   | 2.5         | 145    | 1 OM25_HUMAN | P57105 homo sapien |
| 30         | 173.5 | 2.5         | 569    | 1 AFB3_RAT   | Q70248 rattus norv |
| 31         | 173.5 | 2.5         | 750    | 1 AFB2_RAT   | Q33431 rattus norv |
| 32         | 173   | 2.5         | 2468   | 1 MAPB_HUMAN | P46821 homo sapien |
| 33         | 173   | 2.5         | 3358   | 1 PCGV_MOUSE | Q62059 mus musculu |

|    |       |     |      |              |                     |
|----|-------|-----|------|--------------|---------------------|
| 34 | 172   | 2.5 | 2132 | 1 PCGA_MOUSE | Q61282 mus musculu  |
| 35 | 170   | 2.4 | 736  | 1 DTG2_MOUSE | Q60838 mus musculu  |
| 36 | 169.5 | 2.4 | 2842 | 1 AFB_RAT    | P70478 rattus norv  |
| 37 | 168.5 | 2.4 | 2842 | 1 AFB_MOUSE  | Q61315 mus musculu  |
| 38 | 168   | 2.4 | 1845 | 1 Z236_HUMAN | Q94136 homo sapien  |
| 39 | 167   | 2.4 | 1608 | 1 HLYA_SERMA | P15320 seretelia ma |
| 40 | 164.5 | 2.4 | 2314 | 1 PTP2_HUMAN | P23471 homo sapien  |
| 41 | 164   | 2.3 | 3396 | 1 PCGV_HUMAN | P15611 homo sapien  |
| 42 | 163   | 2.3 | 2843 | 1 AFB_HUMAN  | P25054 homo sapien  |
| 43 | 161.5 | 2.3 | 3172 | 1 ERT3_SACER | Q03133 saccharopol  |
| 44 | 161   | 2.3 | 1577 | 1 HLYA_PROMI | P16466 proteus mir  |
| 45 | 160.5 | 2.3 | 1781 | 1 AKAC_HUMAN | Q02952 homo sapien  |

## ALIGNMENTS

RESULT 1

ID PTND\_HUMAN STANDARD. PRT: 2485 AA

AC Q12923; Q15263; Q16826; Q15264; Q15265; Q15159;

DT 01-OCT-2000 (Rel. 40, Created)

DT 01-OCT-2000 (Rel. 40, Last sequence update)

DT 01-OCT-2000 (Rel. 40, Last annotation update)

DE PROTEIN-TYROSINE PHOSPHATASE, NON-RECEPTOR TYPE 13 (EC 3.1.3.48)

DE (PROTEIN-TYROSINE PHOSPHATASE 1E) (PTP-1E) (PTP-BAS) (PROTEIN-TYROSINE PHOSPHATASE PTP1) (FAS-ASSOCIATED PROTEIN-TYROSINE PHOSPHATASE 1) (FAP-1).

DE PTPN13 OR PTP1E OR PTP1 OR PNP1.

GN Homo sapiens (Human).

OS Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

OX NCBI-TaxID:9606;

RP [1]

RP SEQUENCE FROM N.A.

RC TISSUE-Breast carcinoma;

RX MEDLINE-94350988; PubMed-8071359;

RA Bayville D, Ahmad S, Stocco R, Shen S.-H.;

RT "A novel protein-tyrosine phosphatase with homology to both the cytoskeletal proteins of the band 4.1 family and junction-associated guanylate kinases.";

RT J. Biol. Chem. 269:22320-22327(1994).

RL [2]

RL SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.

RC TISSUE-Leukemia;

RX MEDLINE-94116679; PubMed-8287977;

RA Maekawa K., Imagawa N., Nagamatsu M., Harada S.;

RT "Molecular cloning of a novel protein-tyrosine phosphatase containing a membrane-binding domain and GlGF repeats.";

RL FEBS Lett. 337:200-206(1994).

RL [3]

RL SEQUENCE FROM N.A.

RC TISSUE-Fibroblast;

RX MEDLINE-95014139; PubMed-7929060;

RA Sares J., Claesson-Welsh L., Heldin C.-H., Gonen L.J.;

RT "Cloning and characterization of PTP1, a protein tyrosine phosphatase with similarities to cytoskeletal-associated proteins.";

RL J. Biol. Chem. 269:24082-24089(1994).

RL [4]

RL SEQUENCE OF 1216-2490 FROM N.A.

RC TISSUE-Pancreas;

RX Wang H.Y.;

RT Submitted (JUN-1994) to the EMBL/Genbank/DBS databases.

RL [5]

RL STRUCTURE BY NMR OF 1361-1456.

RX MEDLINE-20170882; PubMed-10704206;

RA Kozlov G., Gehring K., Ekkel I.;

RT "Solution structure of the PDZ domain from human phosphatase hPTP1E and its interactions with C-terminal peptides from the Fas receptor.";

RL Biochemistry 39:2572-2580(2000).

CC -1- FUNCTION: BINDS TO A NEGATIVE REGULATORY DOMAIN IN FAS THAT INHIBITS FAS-INDUCED APOPTOSIS.

[illegible]

OY 1213 GDRIVTIGTSTEGMTHQAVNLKNGSSIEQMVOVAGGVSVTHGHOE-PASSLSFT 1271  
 DB 1831 GRLKRVNDTVNTMTHTDAVNLRAASTVNL-----VIGRVLELPRIPML--- 1877  
 OY 1272 GLTSTSTFODDGLPPQCKSTLERGPDPGISTVGYGSPHGLDLYVKTVPKAAASD 1331  
 DB 1878 -----PHLLPDTITLCKKEELGFSLGGHSLY--QVYVLSIDIMPRVALE 1922  
 OY 1332 GRLKRGDOIIAIVNGOSLSECVTHEEA-----VAIKRTGVTYLM 1370  
 DB 1923 GMLQLDLVHYVNGVSTQGMTELEVRALDMSLPISVLKATRNPLPV 1970  
 RESULT 2  
 DLG2\_RAT 2  
 ID DLG2\_RAT STANDARD: PRF: 852 AA.  
 AC Q63622: Q62939; P70548;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 01-OCT-2000 (Rel. 40, Last annotation update)  
 DE CHANNEL ASSOCIATED PROTEIN OF SYNAPSE-110 (CHAPSIN-110) (SYNAPTIC DENSITY PROTEIN PSD-93) (DISCS, LARGE HOMOLOG 2).  
 GN DLG2.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_Taxid-10116;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-96310881; PubMed-8755482;  
 RA Kim E., Cho K.O., Rothschild A., Sheng M.;  
 RT "Heteromultimerization and NMDA receptor-clustering activity of  
 RT Chapsyn-110, a member of the PSD-95 family of proteins";  
 RL Neuron 17:103-113(1996).  
 RN (2)  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-96193770; PubMed-8625413;  
 RA Breman J.E., Chao D.S., Gee S.H., McGee A.W., Craven S.E.,  
 RA Santillano D.R., Wu Z., Huang F., Xia H., Peters M.F.,  
 RA Froehner S.C., Bredt D.S.;  
 RT "Interaction of nitric oxide synthase with the postsynaptic density protein PSD-95 and alpha1-syntrophin mediated by PDZ domains.";  
 RL Cell 84:757-767(1996).  
 RN (3)  
 RP SEQUENCE FROM N.A.  
 RA Irie M., Hata Y., Takai Y.;  
 RL Submitted (Sep-1996) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: INTERACTS WITH THE CYTOPLASMIC TAIL OF NMDA RECEPTOR SUBUNIT AS WELL AS POTASSIUM CHANNELS.  
 CC -1- SIMILARITY: CONTAINS 3 PDZ/DRH DOMAINS.  
 CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.  
 CC -1- SIMILARITY: BELONGS TO THE MAGUK FAMILY OF CELL JUNCTION PROTEINS.  
 CC -1- SIMILARITY: BELONGS TO THE MAGUK FAMILY OF CELL JUNCTION PROTEINS.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See [http://www.isb-sib.ch/announce/or\\_send\\_an\\_email\\_to\\_license@isb-sib.ch](http://www.isb-sib.ch/announce/or_send_an_email_to_license@isb-sib.ch)).  
 CC or send an email to license@isb-sib.ch).  
 CC EMBL: U49049; AAB53243.1;  
 DR EMBL: U50717; AAC52643.1;  
 DR EMBL: U53368; ABA48562.1;  
 DR HSP: Q12959; IPDR.  
 DR InterPro: IPR000619;  
 DR InterPro: IPR001452;  
 DR InterPro: IPR001478;  
 DR Pfam: PF00625; Guanylate\_kin; 1.  
 DR Pfam: PF00595; PDZ; 3.  
 DR Pfam: PF00018; SH3; 1.  
 DR PROSITE: PS00856; GUANYLATE\_KINASE\_1; 1.

DR PROSITE: PS0052; GUANYLATE\_KINASE\_2; 1.  
 DR PROSITE: PS0006; PDZ; 3.  
 DR PROSITE: PS0002; SH3; 1.  
 KM SH3 domain; Repeat.  
 FT DOMAIN 98 PDZ 1.  
 FT DOMAIN 193 PDZ 2.  
 FT DOMAIN 421 PDZ 3.  
 FT DOMAIN 536 SH3.  
 FT DOMAIN 662 SH3.  
 FT DOMAIN 852 SH3.  
 FT DOMAIN 181 SH3.  
 FT DOMAIN 228 SH3.  
 FT DOMAIN 326 SH3.  
 FT DOMAIN 339 SH3.  
 FT DOMAIN 450 SH3.  
 FT DOMAIN 464 SH3.  
 FT DOMAIN 474 SH3.  
 FT DOMAIN 476 SH3.  
 FT DOMAIN 478 SH3.  
 FT DOMAIN 484 SH3.  
 FT DOMAIN 506 SH3.  
 FT DOMAIN 569 SH3.  
 FT DOMAIN 586 SH3.  
 FT DOMAIN 626 SH3.  
 FT DOMAIN 639 SH3.  
 FT DOMAIN 726 SH3.  
 FT DOMAIN 733 SH3.  
 FT DOMAIN 749 SH3.  
 FT DOMAIN 756 SH3.  
 FT DOMAIN 792 SH3.  
 FT DOMAIN 794 SH3.  
 SO SEQUENCE 852 AA; 94934 MW; FBD41AAB9CF5B09 CAC64;  
 Query Match 5.8%; Score 405.5; DB 1; Length 852;  
 Best Local Similarity 28.4%; Pred. No. 4.6e-14;  
 Matches 155; Conservative 67; Mismatches 213; Indels 111; Gaps 22;  
 OY 866 EMPDQAVNSAAGASG-----EKKNS-----QSLMPOSSGPEESTIRNRS 909  
 DB 22 DGPDHSLRLTHEYRGPELVHSEKNSQIENVHGYLSHSPKASPAV-IIVNTDT 80  
 OY 910 SSTPAIFASDPATCPPIPCETITIE-----ISKRTGLGLSTVGGSDT-LIG---AFIIN 960  
 DB 81 LD-----TIYVNTLEIYEFEETTLERNGSGLGSFISAGTDNPHIDDDGIFIT 130  
 OY 961 EYERGAACKGRILNAGDOILEVNGIDLRKATHEAIVNLYKOTPORVRLTYRDEAPYKE 1020  
 DB 131 KIIPGGAALDGRILRVNDCILRVNEVDVSEVSHKAVELAKGASIVRLYV-RRRRPILE 189  
 OY 1021 EYVCDTLTLEQKKPKGKGLSIVK-----RNDTGVVSVIIVGGIADPGRILOGDO 1074  
 DB 190 TVV-----EIKLKGK-PKGLGFSIAGVGNQHIIPGDSIIVTKIIDGGAOKKGRLOVGR 244  
 OY 1075 ILLVNGEDVRNASEVAVALNCSLQTVTLEVGRIKA-----GP---FHSERRPSQTS 1124  
 DB 245 LLMVNNYSLEEVTHEBAVAIILKNTSDVYLVKKGKPTIYMTDPPGPDITHSYSPMEWH 304  
 OY 1125 QVSEGLSFTPLPGSSITSESSLESSKKNALASELQGRITVEAKKGPDISIGISIAGV 1184  
 DB 305 LLS-GNNGLEKTSLLPISPRYSPIPKMLV-BDDYTRPE---PYSTVKNLCCKP 358  
 OY 1185 GSPLDGVPIFIIMMHPITGVAOQKLRVGDRIIVTIGTSTEGMTHQAVNLKNGSSIE 1244  
 DB 359 ASPRKSPV-----ECDSNF-----LSTPYPHH 383  
 OY 1245 MOVVAGGVSVTGHQEPASSLSFTGLTSTSTFODDGLPPQCKSTLERGPDPGLGFSI 1304  
 DB 384 LCLLPDSDMTSHQSTATRQSPVTLQRAISLE-----GEP--RKVYLHKSGTGLGFI 435  
 OY 1305 VCGYSGPHDLPITYKTVAKKAAASDGLKRGDOIIAIVNGOSLSECVTHEEAVALIKRK 1364  
 DB 436 VGG---EDGE-GIFVSFIILAGPADLSELQRGDOILSVNGIDLRGASHDEQAAALKGAG 491



|                       |                  |                    |             |             |
|-----------------------|------------------|--------------------|-------------|-------------|
| Query Match           | 5.7%             | Score 397.5;       | DB 1,       | Length 870; |
| Best Local Similarity | 28.1%            | Pred. No. 1.2e-13; |             |             |
| Matches 154;          | Conservative 69; | Mismatches 205;    | Indels 121; | Gaps 23;    |

```

Db 433 ENIYGG---EDGE-GIFP5FLIAGGPNLDELQRGDQIILSVNGIDLRGASHEDQAAALAK 488
OY 1362 RTKGTVILM 1370
Db 489 GAGQTVIIL 497

RESULT 5
ID DLG3_MOUSE STANDARD; PRT; 849 AA.
AC P70175;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE 01-OCT-2000 (Rel. 40, Last annotation update)
DE PRESYNAPTIC PROTEIN SAPI02 (SINAPSE-ASSOCIATED PROTEIN 102) (DISCS,
DE LARGE HOMOLOG 3).
GN DLG3 OR DLG3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RX NCBI_TaxID=10090;
[1]
RA SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Brain;
RL Kohmura N., Makino S., Yagi T.;
Submitted (Aug-1996) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: INTERACTS WITH THE CYTOPLASMIC TAIL OF THE NMDA RECEPTOR
CC SUBUNIT NR2B.
CC -1- SIMILARITY: CONTAINS 3 PDZ/DHR DOMAINS.
CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
CC -1- SIMILARITY: CONTAINS A GUANYLATE KINASE-LIKE DOMAIN.
CC -1- SIMILARITY: BELONGS TO THE MAGOK FAMILY OF CELL JUNCTION PROTEINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: D87117; BAA13249.1; -.
CC HSSP: A12859; IPDR.
CC MGD: MGI:1886886; Dlg3.
CC InterPro: IPR000619; -.
CC InterPro: IPR001452; -.
CC InterPro: IPR001478; -.
CC Pfam: PF006625; Guanylate_kin; 1.
CC Pfam: PF00595; PDZ; 3.
CC Pfam: PF00018; SH3; 1.
CC PROSITE: PS00856; GUANYLATE_KINASE_1; 1.
CC PROSITE: PS50052; GUANYLATE_KINASE_2; 1.
CC PROSITE: PS50106; PDZ; 3.
CC PROSITE: PS50002; SH3; 1.
CC SH3 domain; Repeat.
CC K1 DOMAIN 149 235 PDZ 1.
CC FT DOMAIN 244 330 PDZ 2.
CC FT DOMAIN 404 484 PDZ 3.
CC FT DOMAIN 519 589 SH3.
CC FT DOMAIN 659 849 GUANYLATE_KINASE.
CC SO SEQUENCE 849 AA; 93482 MW; EF3EF2D513538BE CRC64;

Query Match 5.6%; Score 393; DB 1; Length 849;
Best Local Similarity 24.7%; Pred. No. 2,le-13;
Matches 139; Conservative 73; Mismatches 173; Indels 178; Gaps 18;

OY 868 PDSQAVPAAGAAAGSEKKNSSQSLNVPQSGSPDEPS---IRNTSR-----SST 912
Db 36 PDPEYSGGNGASGSSGYGVSSQTLPSQAGARTPTRTAKKLIPYTGADVGPVPKPVPGKST 95
OY 913 PAIRFSDP-----ATC-----PIIPGCT-----TLEISK 938

```

```

Dy 96 PKLNGSGPMPRECTCTNRDWEQASPAFLVNPDALEPSLVNGSDGMKEEYEVLERG 155
Oy 939 RTGGLSTVGGSDTL-----LGAFIHEVYEGAAKDGRLMAGDQILEVNGIDLRKATH 993
Dy 156 NSGLGFSIAGIDPHVDPDPGIFIT-KIIPGGAAMDGRGVNDVLRVNEVDVSEVYH 214
Oy 994 DEAINVLKOTPPQVRLTYLDEAPYKEEVEDTLTIELOKKPKGKGLSTVYK-----R 1047
Dy 215 SRAVEALKKAGPVRLVLRVRRQRPPE-----TIMEVNLKGP-KGLGFSIAGIGNGHIP 268
Oy 1048 NDTGVFVSDIYKGIADPDGRLIOGDQILLVNGEDVRNASOEAVALKCSLGTVTLEVG 1107
Dy 269 GDNISYITKIIIEGGAOKDGRLOIGDRLLAVNNTNLODVHREEVAVASLKTNSDMVYLKVA 328
Oy 1108 RIKGPPHSEKRPSTQSVSGSLSTFPLSGSTSESSSKKNAALSEIQLRTVE 1167
Dy 329 --KPGSLH-----LNDMYAPPDYASTFTALADNH----- 355
Oy 1168 MKKGPTDSLGISIAGVSPGLGVPIFIAHMHPTVGAQOTKLRVDRIVTIGTSTEGM 1227
Dy 356 -----I 356
Oy 1228 THFOAVNLKKNASISIMQVYAGDVSVYTGHHOEPASSLSFTGLSTSFODDLGPQ 1287
Dy 357 SHNSLSLYL-----GAVESKVT-----YAPAPQVPTPRYSPIPRHMLAEEDF-TRE 401
Oy 1288 CKSTTLERGPDLGFSIYGVGSPHGLPIYKTVFAKGAASEGRRLKRGDQIIAVNGOS 1347
Dy 402 PKRIILKKGSTGLGFNTVGG---EDGE-GIFVSTILAGPADLSGELRGDRILSVNGVN 457
Oy 1348 LEGVTHEEVAAILKRTKGTVTLM 1370
Dy 458 LRNATHEQAAALAKRAGOSVTIV 480

```

RESULT 6

ID DUG3\_RAT STANDARD; PRT: 849 AA.

AC 062936; P70547;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 01-OCT-2000 (Rel. 40, Last annotation update)  
 DE PRESTNAPIC PROTEIN SAP102 (SYNAPSE-ASSOCIATED PROTEIN 102) (PSD-95/SAP90 RELATED PROTEIN 1) (DISCS, LARGE HOMOLOG 3).  
 GN DUG3.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclerogamathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A. (LONG FORM).  
 RC TISSUE=Brain;  
 RX MEDLINE=96374358; Pubmed=8780649;  
 RA Mueller B.M., Kistner U., Kindler S., Chung W.J., Kuhlendahl S., Fensler S.D., Lau U.F., Voh R.W., Hugenir R.L., Gundelfinger E.D., Garner C.C.;  
 RA "SAP102, a novel postsynaptic protein that interacts with NMDA receptor complexes in vivo.";  
 RT Neuron 17:255-265(1996).  
 RN [2]  
 RP SEQUENCE FROM N.A. (SHORT FORM).  
 RA Title M., Hata Y., Takai Y.;  
 RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.  
 CC -FUNCTION: INTERACTS WITH THE CYTOPLASMIC TAIL OF THE NMDA RECEPTOR SUBUNIT NR2B.  
 CC -SIMILARITY: CONTAINS 3 PDZ/DHR DOMAINS.  
 CC -SIMILARITY: CONTAINS 1 SH3 DOMAIN.  
 CC -SIMILARITY: CONTAINS A GUANYLATE KINASE-LIKE DOMAIN.  
 CC -SIMILARITY: BELONGS TO THE MAGOK FAMILY OF CELL JUNCTION PROTEINS.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.ebi.ac.uk/announcements> or send an email to [license@ebi.ac.uk](mailto:license@ebi.ac.uk)).

CC EMBL; U50147; AAA93031.1; -;  
 CC EMBL; U53367; AAB48561.1; -;  
 CC HSSP; Q12959; IPDR.  
 CC InterPro; IPR000619; -;  
 CC InterPro; IPR001452; -;  
 CC InterPro; IPR001478; -;  
 CC Pfam; PF00625; Guanylate\_kin.1.  
 CC Pfam; PF00595; PDZ.3.  
 CC Pfam; PF00018; SH3.1.  
 CC PROSITE; PS00856; GUANYLATE\_KINASE\_1; 1.  
 CC PROSITE; PS0052; GUANYLATE\_KINASE\_2; 1.  
 CC PROSITE; PS50106; PDZ.3.  
 CC PROSITE; PS50002; SH3.1.  
 CC SH3 domain; Repeat; Alternative splicing.  
 CC PDZ 1.  
 CC PDZ 2.  
 CC PDZ 3.  
 CC SH3.  
 CC GUANYLATE\_KINASE.  
 CC MISSING (IN SHORT ISOFORM).  
 CC VARSPLIC 627 640  
 CC SEQUENCE 849 AA; 93539 MW; 34DA9C46C7BB6DB CRC64;

Query Match 5.68; Score 391; DB 1; Length 849;  
 Best Local Similarity 24.58; Pred. No. 2,7e-13;  
 Matches 138; Conservative 75; Mismatches 172; Indels 178; Gaps 18;

```

Oy 868 PDSQAVPSAAGAAAGKSSQSLMPQSGPEPES---IKNTSR-----SST 912
Dy 36 PDYGPSPGNGKASSGYSQTLPSQAGANTPRTRAKILPTGRDVGPPRPKPKNT 95
Oy 913 PAIPASDPA-----TC-----PIIPGCT-----TIEISKG 938
Dy 96 PKLNGSGPMPRECTCTNRDWEQASPAFLVNPDALEPSLVNGSDGMKEEYEVLERG 155
Oy 939 RTGGLSTVGGSDTL-----LGAFIHEVYEGAAKDGRLMAGDQILEVNGIDLRKATH 993
Dy 156 NSGLGFSIAGIDPHVDPDPGIFIT-KIIPGGAAMDGRGVNDVLRVNEVDVSEVYH 214
Oy 994 DEAINVLKOTPPQVRLTYLDEAPYKEEVEDTLTIELOKKPKGKGLSTVYK-----R 1047
Dy 215 SRAVEALKKAGPVRLVLRVRRQRPPE-----TIMEVNLKGP-KGLGFSIAGIGNGHIP 268
Oy 1048 NDTGVFVSDIYKGIADPDGRLIOGDQILLVNGEDVRNASOEAVALKCSLGTVTLEVG 1107
Dy 269 GDNISYITKIIIEGGAOKDGRLOIGDRLLAVNNTNLODVHREEVAVASLKTNSDMVYLKVA 328
Oy 1108 RIKGPPHSEKRPSTQSVSGSLSTFPLSGSTSESSSKKNAALSEIQLRTVE 1167
Dy 329 --KPGSLH-----LNDMYAPPDYASTFTALADNH----- 355
Oy 1168 MKKGPTDSLGISIAGVSPGLGVPIFIAHMHPTVGAQOTKLRVDRIVTIGTSTEGM 1227
Dy 356 -----I 356
Oy 1228 THFOAVNLKKNASISIMQVYAGDVSVYTGHHOEPASSLSFTGLSTSFODDLGPQ 1287
Dy 357 SHNSLSLYL-----GAVESKVT-----YAPAPQVPTPRYSPIPRHMLAEEDF-TRE 401
Oy 1288 CKSTTLERGPDLGFSIYGVGSPHGLPIYKTVFAKGAASEGRRLKRGDQIIAVNGOS 1347
Dy 402 PKRIILKKGSTGLGFNTVGG---EDGE-GIFVSTILAGPADLSGELRGDRILSVNGVN 457
Oy 1348 LEGVTHEEVAAILKRTKGTVTLM 1370
Dy 458 LRNATHEQAAALAKRAGOSVTIV 480

```



```

RESULT 7
ID      DLG4_HUMAN          STANDARD:          PRT:          767 AA.
AC      P78352; Q92941; Q9UKR8;
DT      01-NOV-1997 (Rel. 35; Created)
DT      30-MAY-2000 (Rel. 39; Last sequence update)
DT      01-OCT-2000 (Rel. 40; Last annotation update)
DE      PRESYNAPTIC DENSITY PROTEIN 95 (PSD-95) (DISCS, LARGE HOMOLOG 4).
GN      DLG4 OR PSD95.
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX      NCBI_Taxid=9606;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      TISSUE=Mammary gland;
RC      MEDLINE=97432822; PubMed=9286702;
RA      Stachakis D.G., Hoover K.B., You Z., Bryant P.J.;
RT      "Human postsynaptic density-95 (PSD95): location of the gene (DLG4)
RT      and possible function in nonneural as well as in neural tissues.";
RL      Genomics 44:71-82(1997).
RN      [2]
RP      REVISIONS.
RC      TISSUE=Mammary gland;
RC      Stachakis D.G., Hoover K.H., You Z., Bryant P.J.;
RL      Submitted (Jul-1998) to the EMBL/Genbank/DBJ databases.
RN      [3]
RP      SEQUENCE FROM N.A.
RA      Stachakis D.G., Udar N., Sandgren O., Andreasson S., Bryant P.J.,
RA      Small K., Forsman-Semb K.;
RT      "Genomic organization of human DLG4, the gene encoding post-synaptic
RT      density 95 (PSD95).";
RL      J. Neurochem. 73:0-0(1999).
RN      [4]
RP      SEQUENCE OF 81-401 FROM N.A.
RC      TISSUE=Brain;
RC      Breman J.E., Bredt D.S., Parkinson J.F., Manzana W.P., McClary J.A.;
RL      Submitted (Aug-1996) to the EMBL/Genbank/DBJ databases.
CC      -1- FUNCTION: INTERACTS WITH THE CYTOPLASMIC TAIL OF NMDA RECEPTOR
CC      SUBUNITS. MAY BE INVOLVED IN SYNAPTONEGENESIS.
CC      -1- SUBCELLULAR LOCATION: CONCENTRATED AT SYNAPTIC JUNCTIONS PRIMARILY
CC      ON THE PRESYNAPTIC SIDE (WAS ORIGINALLY THOUGHT TO BE
CC      POSTSYNAPTIC).
CC      -1- TISSUE SPECIFICITY: PRESYNAPTIC DENSITY FRACTION OF BRAIN.
CC      -1- SIMILARITY: CONTAINS 3 PDZ/DHR DOMAINS.
CC      -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
CC      -1- SIMILARITY: CONTAINS A GUANYLATE KINASE-LIKE DOMAIN.
CC      -1- SIMILARITY: BELONGS TO THE MAGUK FAMILY OF CELL JUNCTION PROTEINS.
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@sib-sib.ch).
CC      -----
CC      EMBL: U83192; AAC52113.1;
CC      EMBL: AF156495; RAD56173.1;
CC      EMBL: U68138; AAB07736.1;
CC      HSSP: 012959; 1PDR.
CC      MIM: 602887;
CC      InterPro: IPR000619;
CC      InterPro: IPR001452;
CC      InterPro: IPR001478;
CC      Pfam: PF00625; Guanylate_kin; 1.
CC      Pfam: PF00595; PDZ; 3.
CC      Pfam: PF00018; SH3; 1.
CC      PROSITE: PS00856; GUANYLATE_KINASE_1; 1.
CC      PROSITE: PS50052; GUANYLATE_KINASE_2; 1.
CC      PROSITE: PS50106; PDZ; 3.
CC      PROSITE: PS50002; SH3; 1.
CC      SH3 domain; Repeat.

```

```

FT      DOMAIN          108      194      PDZ 1.
FT      DOMAIN          203      289      PDZ 2.
FT      DOMAIN          356      436      PDZ 3.
FT      DOMAIN          471      541      SH3.
FT      DOMAIN          577      767      GUANYLATE KINASE.
FT      CONFLICT        46      46      E -> V (IN REF. 3).
FT      CONFLICT        81      83      VIV -> EFR (IN REF. 4).
FT      CONFLICT        399      401      GDO -> AGI (IN REF. 4).
SQ      SEQUENCE        767 AA; 85429 MW; BE101915965B2D8 CRC64;

Query Match          5.5%; Score 384; DB 1; Length 767;
Best Local Similarity 25.8%; Pred. No. 5.3e-13;
Matches 131; Conservative 76; Mismatches 153; Indels 148; Gaps 18;

QY      886 NSQSILMPQSSP-----EPESIRNTSRSTPAIRASDPATCP-----IIPGCTTIE--- 934
DB      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      50 SESQKRYODEDPTPLPESHPAHLPNQA-NSPPVYVTDPLPAGVELQVNGTEGEMEYEE 108
DB      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      935 --ISKRTGLGISVGGSPD-LLG---AFIHEVVEGAACRGRMAGDOILEVNGIDL 988
DB      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      109 ILEKNSGLGFSIAGTDNPHIGDDPSIFTKIIPGGAADGRLRVNDLSILEVNEVDV 168
DB      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      989 KRATHEAIVNLKQTPQRYRLTLYRDEAFYKEEYCDTLIELOKKPKGLGISYVK-- 1046
DB      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      169 REYTHSAVAEALKEAGSIYRLVYMRKPPAEK-----VMEIKLIKRP-KGLGFSINGVG 222
DB      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      1047 ---RNDTGVFSDIYKGIADPDGRLIOGDOILLVNGEDVNNASQEAVALLKSLGTV 1102
DB      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      223 NQHIPGDNISYVTKILEGGAHRKDRLQIGDKILANVSGLEPVMEDAVALKNTYDVV 282
DB      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      1103 TLEVGRIRKAPFSESRPQTSQVSEGLSFPPLSGSSTSESSSKKNALASEIIG 1162
DB      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      283 YLKVA-----KPS-NATLSD---STAPPITTSYSHLDNEISHS----- 319
DB      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      1163 LRTVENKKGPTDLSIGISIAGVGSPGLGDVPIFAMHPGVAAQOTOKLRVGRITYTCT 1222
DB      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      320 -----YLGDTYPT----- 327
DB      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      1223 STEGMTHQAVNLKKNASGSIEMQVAGDVSYVGHQHPASSLSFTGLSTSTFPOD 1282
DB      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      328 ---AMPT-----SPRRYSPVAKDLIGED 349
DB      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      1283 LGPPKSTILERPGDGLFSIYGGYSPHGLPIYKTVFAKGAASEGRKRGDOITA 1342
DB      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      350 I-PREPRRIIVHNGSTGLFNTVGG---EDGE-GIFSTIAGPADLSGELRKQJILS 404
DB      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      1343 VNGQSLEGVTHEEVAAILKRTGVTIM 1370
DB      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      405 VNGVDLRNASHQDAIALKNAQOTVIT 432
DB      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 8
ID      DLG4_RAT          STANDARD:          PRT:          724 AA.
AC      P31016; P97631;
DT      01-JUL-1993 (Rel. 26; Created)
DT      01-JUL-1993 (Rel. 26; Last sequence update)
DT      01-OCT-2000 (Rel. 40; Last annotation update)
DE      PRESYNAPTIC DENSITY PROTEIN 95 (PSD-95) (PRESYNAPTIC PROTEIN SAP90)
DE      (SYNAPTIC-ASSOCIATED PROTEIN 90) (DISCS, LARGE HOMOLOG 4).
GN      DLG4 OR PSD95.
OS      Rattus norvegicus (Rat).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX      NCBI_Taxid=10116;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=SPRAGUE-DAWLEY; TISSUE=Brain;
RC      MEDLINE=93040233; PubMed=1419001;
RA      Cho K.-O., Hunt C.A., Kennedy M.B.;
RT      "The rat brain postsynaptic density fraction contains a homolog of
RT      the Drosophila discs-large tumor suppressor protein.";

```

```

NL Neuron 9:929-942(1992).
[2]
RN
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=Brain;
RX MEDLINE=93186749; PubMed=7680343;
RA Kistner U., Wenzel B.M., Vehl R.M., Cases-Langhoff C., Garner A.M.,
RA Appelbauer U., Voss B., Gundelfinger E.D., Garner C.C.;
RA "SAP90, a rat presynaptic protein related to the product of the
RA Drosophila tumor suppressor gene dlg-A.";
RA J. Biol. Chem. 268:4580-4583(1993).
RN
RN
RN
RP SEQUENCE OF 566-625 FROM N.A.
RA STRAIN=WISTAR K170; TISSUE=vascular smooth muscle;
RA Adams L.D., Werny I., Schwartz S.M.;
RA Submitted (NOV-1996) to the EMBL/Genbank/DBJ databases.
[4]
RN
RP X-RAY CRYSTALLOGRAPHY (1.82 ANGSTROMS) OF 302-402.
RX MEDLINE=96270509; PubMed=8674113;
RA Doyle D.A., Lee A., Lewis J., Kim E., Sheng M., Mackinnon R.;
RA "Crystall structures of a complexed and peptide-free membrane protein-
RA binding domain: molecular basis of peptide recognition by PDZ.";
RA Cell 85:1067-1076(1996).
CC
CC -1 FUNCTION: INTERACTS WITH THE CYTOPLASMIC TAIL OF NMDA RECEPTOR
CC SUBUNTS. MAY BE INVOLVED IN SYNAPTOGENESIS.
CC -1 SUBCELLULAR LOCATION: CONCENTRATED AT SYNAPTIC JUNCTIONS PRIMARILY
CC ON THE PRESYNAPTIC SIDE (WAS ORIGINALLY THOUGHT TO BE
CC POSTSYNAPTIC).
CC
CC -1 TISSUE SPECIFICITY: PRESYNAPTIC DENSITY FRACTION OF BRAIN.
CC -1 SIMILARITY: CONTAINS 3 PDZ/DHR DOMAINS.
CC -1 SIMILARITY: CONTAINS 1 SH3 DOMAIN.
CC -1 SIMILARITY: CONTAINS A GUANYLATE KINASE-LIKE DOMAIN.
CC -1 SIMILARITY: BELONGS TO THE MAGUK FAMILY OF CELL JUNCTION PROTEINS.
CC
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M96853; AAA41971.1; -
DR EMBL; X66474; CAA47103.1; -
DR EMBL; U77090; AAB38270.1; -
DR PIR; S26407; S26407.
DR PIR; JH0800; JH0800.
DR PDB; 1BE9; 2I-OCT-98.
DR PDB; 1BE9; 2I-OCT-98.
DR InterPro; IPR000619; -
DR InterPro; IPR001452; -
DR InterPro; IPR001478; -
DR Pfam; PF00625; Guanylate_kin; 1.
DR Pfam; PF00595; PDZ; 3.
DR Pfam; PF00018; SH3; 1.
DR PROSITE; PS00856; GUANYLATE_KINASE_1; 1.
DR PROSITE; PS50052; GUANYLATE_KINASE_2; 1.
DR PROSITE; PS50106; PDZ; 3.
DR PROSITE; PS50002; SH3; 1.
KW SH3 domain; Repeat; 3D-structure.
FT
FT DOMAIN 65 151 PDZ 1.
FT DOMAIN 160 246 PDZ 2.
FT DOMAIN 313 393 PDZ 3.
FT DOMAIN 428 498 SH3.
FT DOMAIN 534 724 GUANYLATE KINASE.
FT CONFLICT 61 61 M -> L (IN REF. 2).
FT CONFLICT 78 78 S -> T (IN REF. 2).
FT CONFLICT 177 182 S -> T (IN REF. 2).
FT CONFLICT 200 200 GAGMOH -> ALGTSTI (IN REF. 2).
FT CONFLICT 254 254 S -> T (IN REF. 2).
FT CONFLICT 539 555 ILGPTKQKANDDLSEF -> ISLDPKPTVPTMIFSPSS
FT (IN REF. 2).
FT CONFLICT 623 625 GKH -> RDO (IN REF. 3).
FT SEQUENCE 724 AA; 80465 MW; 7922D4EBE0F9AD85 CRC64;

```

Query Match Summary: 5.5%, score 383; DB 1; Length 724;  
Best Local Similarity: 26.2%; Pred. No. 5.5e-13;  
Matches: 128; Conservative: 74; Mismatches: 143; Indels: 144; Gaps: 17;

|    |      |                             |  |                     |
|----|------|-----------------------------|--|---------------------|
| OY | 901  | PESINSTRSSSTPAIFASDPTCP---- | IIPCEFTIE-----ISKRTGLSLGVGSD                 | 951                 |
| Db | 26   | PAHLNQA--NSPPVIVNTDLAPYEI   | VQNVTEGEEMEETTLERGNLSGFISHTD                 | 84                  |
| OY | 952  | T-LIG--AFITHEVEYGACCKDR     | LWAGDOILEVNGIDLKATNDEINVLRQTPOKV             | 1007                |
| Db | 85   | NPHIDDPSIFTTKIIRPGAAAO      | DRLKYNDISILVEYNDAEVTHSAVAELKAGSIV            | 144                 |
| OY | 1008 | RLLTVRDEAPRKKEEVCOTLT       | ELOKKPKGIGLSIVK-----RNDTGVFVSIVKGG           | 1061                |
| Db | 145  | RLYVMRRKPRAEK-----          | VMETIKLIKRP-KGLGSFIAGVGQNII                  | PCDNSIVYTKIIECG 198 |
| OY | 1062 | IADPRGRLOGDOIILLVGED        | EDVRNMSOEVAVALMLCSLGTYILEVGR                 | KAGFHSESRPS 1121    |
| Db | 139  | AANKRGRIODGIKLILAUV         | SGLVEDYMHDDAVAALKNYTDVVYLKA-                 | -RPS 247            |
| OY | 1122 | OTSQVSEGLSFEPTPLSGSS        | TSSESLESSKKNALASEIOGLRTVEMKGPDSIGISA         | 1181                |
| Db | 248  | NAYLSD----SYARPDITTSYS      | QLHDEMISHS-----                              | 276                 |
| OY | 1182 | GCVSPCLDDVPFLFMHP           | TPGYAOTAOKLRNGDRITVTTCISTEBSMTHTQAVNLKNMSG   | 1241                |
| Db | 277  | -----YLDTPPT-----           | --MTPP-----                                  | 289                 |
| OY | 1242 | SIEMQVAVGDVSVYTGHN          | HOEPASSLSFGTLSTISIFODLGRPROCKSITYLERPDGLG    | 1301                |
| Db | 290  | -----SPKRYSPAKOLLGEED       | -PREPRIYVHNRSSTGLG                           | 324                 |
| OY | 1302 | FSTVGGYSSPHGDLPIYVK         | YVFAPKCAASEDGRLKGDQOIIAVNGOSLEGVTHEEAVALIK   | 1361                |
| Db | 335  | FNIVGG---EDGE-GIFS          | IFLAGPADLSCGLERRKDQOIIISYNGVDLNRNASHSQAAIALK | 380                 |
| OY | 1362 | RTKGVTLM                    | 1370   |                     |
| Db | 381  | NAGQTVTTI                   | 389  |                     |

RESULT 9  
DLG4\_MOUSE STANDARD: PRT: 724 AA.  
ID DLG4\_MOUSE  
AC Q62108:  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 01-OCT-2000 (Rel. 40, Last annotation update)  
DE PRESYNAPTIC DENSITY PROTEIN 95 (PSD-95) (PRESYNAPTIC PROTEIN SAP90)  
DE (SYNAPSE-ASSOCIATED PROTEIN 90) (DISCS, LARGE HOMOLOG 4).  
GN DLG4 OR DLG4H OR PSD95.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrate; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=DBA/2; TISSUE=Brain;  
RA Kohmura N., Yaagi T.;  
RL Submitted (May-1995) to the EMBL/GenBank/DBJ databases.  
CC -1- FUNCTION: INTERACTS WITH THE CYTOPLASMIC TAIL OF NM2A RECEPTOR SUBUNITS. MAY BE INVOLVED IN SYNAPTOGENESIS (BY SIMILARITY).  
CC -1- SUBCELLULAR LOCATION: CONCENTRATED AT SYNAPTIC JUNCTIONS PRIMARILY ON THE PRESYNAPTIC SIDE (BY SIMILARITY).  
CC -1- SIMILARITY: CONTAINS 3 PDZ/DHR DOMAINS.  
CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.  
CC -1- SIMILARITY: CONTAINS A GUANYLATE KINASE-LIKE DOMAIN.  
CC -1- SIMILARITY: BELONGS TO THE MAGUK FAMILY OF CELL JUNCTION PROTEINS.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation -

CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sdb.ch/announce/>  
 CC or send an email to [license@sdb.ch](mailto:license@sdb.ch)).

DR EMBL: D50621; BAA09297.1; -  
 DR HSP: Q12959; 1PDR.  
 DR MGD: MGI:1277959; DlgH4.  
 DR InterPro: IPR000619; -  
 DR InterPro: IPR001452; -  
 DR InterPro: IPR001478; -  
 DR Pfam: PF00625; Guanylate\_kin; 1.  
 DR Pfam: PF00595; PDZ; 3.  
 DR Pfam: PF00018; SH3; 1.  
 DR PROSITE: PS00856; GUANYLATE\_KINASE\_1; 1.  
 DR PROSITE: PS50052; GUANYLATE\_KINASE\_2; 1.  
 DR PROSITE: PS50106; PDZ; 3.  
 DR PROSITE: PS50002; SH3; 1.  
 DR SH3 domain; Repeat.  
 FT DOMAIN 65 151 PDZ 1.  
 FT DOMAIN 160 246 PDZ 2.  
 FT DOMAIN 313 393 PDZ 3.  
 FT DOMAIN 428 498 SH3.  
 FT DOMAIN 534 724 GUANYLATE\_KINASE.  
 SO SEQUENCE 724 AA; 80472 MW; 7EFEC99E1FF90BA CRC64;

Query Match 5.5%; Score 382; DB 1; Length 724;  
 Best Local Similarity 26.2%; Pred. No. 6.2e-13;  
 Matches 128; Conservative 74; Mismatches 143; Indels 144; Gaps 17;

QY 901 PESIRTSSTSPATPASPATP-----IIPCETTF-----ISKRGGLGISTYGGSD 951  
 DB 26 PAHLPAQA--NSPPVYNTLTLEAPGELOYNGEGMEYEETTLERNGSLGFSIAGTD 84  
 QY 952 T-LG---AFIHEVEEGACGKGRMAGDOILEVNGIDLRKATDEALINVLROTPQWY 1007  
 DB 85 NHIGDPSIFITKIRPGMAADGRLRVNDSTLFEVNEVREVTISAVALKEAGSIY 144  
 QY 1008 RLTLYRDEAPYKEEYVCDLTLELQKPKGKGLGSLIVK-----RNDGVFVSDIVKGG 1061  
 DB 145 RLYVMRRKPPAEK-----IIEIKLIKGP--KGLGFSIAGGVGNQIHIGDINSIYVTKIIEGG 198  
 QY 1062 IADPGRRLQSGQIILLVNEEDVNAQSEVAALLKSLGVTLVEVRIAGPRHSRRS 1121  
 DB 199 AAHKDRLQIGKILAVNVEGVLEDAVNAALKNVYVYLKVA-----KPS 247  
 QY 1122 QTSOVSEGLSFTPLSGSTSESSSKNALASEIQLRTVEMKKGPDSIGISIA 1181  
 DB 248 -NAYLSD-----SYAPPDTTYSQHLDNELSHS----- 276  
 QY 1182 GGVSPLGDPVPIFIAMHPTGVAAGTOKLRVGDRIYTCGTSTEGMTHTQAVNLLKNASG 1241  
 DB 277 -----YLGTDYPT-----AATPT----- 289  
 QY 1242 SIEMOVVAGDVSVYTGHHOFRASSLSFTGLSTSIIFODDLGPRCKSTLERGPDGIC 1301  
 DB 290 -----SPRRYSVPAKDLLGEEDL--PREPRRIYVHRSGTGIG 324  
 QY 1302 FSVIGGSGPHGDLPIYKTVFKAASDEGRKRGDQIIAANGSLLEGVTHEEVAAILK 1361  
 DB 325 FNVVGG---EDGE-GIFISFIILAGPADLSGELRKDQILSVNGVLDLRNASHQAAILK 380  
 QY 1362 RTKGVITLM 1370  
 DB 381 NAGQVTVII 389

RESULT 10  
 DGL1\_DROME STANDARD; PRT; 960 AA.  
 ID DGL1\_DROME  
 AC P31007;

DR 01-JUL-1993 (Rel. 26; Created)  
 DT 01-JUL-1993 (Rel. 26; Last sequence update)  
 DT 30-MAY-2000 (Rel. 39; Last annotation update)  
 DE LETHAL(1)DISCS LARGE-1 TUMOR SUPPRESSOR PROTEIN.  
 GN DGL1.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 ON NCBI\_TaxId=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=91330294; PubMed=1651169;  
 RA Woods D.F., Bryant P.J.;  
 RT "The disc-large tumor suppressor gene of Drosophila encodes a  
 RT guanylate kinase homolog localized at septate junctions.";  
 RL Cell 66:451-464(1991).  
 CC -1- FUNCTION: LOSS OF THIS PROTEIN CAUSES LARGE IMAGINAL DISKS BY  
 CC ALLOWING NEOPLASTIC OVERGROWTH. THE PRESENCE OF A GUANYLATE KINASE  
 CC DOMAIN SUGGESTS INVOLVEMENT IN ADEQUATE CELLULAR ADHESION AS WELL  
 CC AS SIGNAL TRANSDUCTION TO CONTROL CELLULAR PROLIFERATION. THE DGL  
 CC PROTEIN PLAYS A CRITICAL ROLE AT SEPTATE JUNCTIONS IN CELLULAR  
 CC GROWTH CONTROL DURING LARVAL DEVELOPMENT.  
 CC -1- SUBCELLULAR LOCATION: CYTOSKELETON-ASSOCIATED. LOCATED AT THE  
 CC CYTOPLASMIC FACE OF THE MEMBRANE IN THE CELLULAR BLASTODERM AND  
 CC BECOMES ASSOCIATED WITH SEPTATE JUNCTIONS WHICH BEGIN TO FORM  
 CC BETWEEN EPITHELIAL CELLS AT THE TIME OF DORSAL CLOSURE. IN ADULT  
 CC FLIES, LOCATED AT THE APICAL-LATERAL MEMBRANE BOUNDARY OF  
 CC EPITHELIAL CELLS.  
 CC -1- ALTERNATIVE PRODUCTS: LEADS TO DIFFERENT FORMS OF THE PROTEIN;  
 CC THE SEQUENCE SHOWN REFERS TO THE DGL-A PROTEIN, THAT IS MOST  
 CC ABUNDANTLY EXPRESSED.  
 CC -1- TISSUE SPECIFICITY: BROAD SPECIFICITY. ABUNDANTLY EXPRESSED IN  
 CC EPITHELIAL TISSUES. DEVELOPMENTAL STAGE DEPENDENT EXPRESSION.  
 CC -1- SIMILARITY: CONTAINS 3 PDZ/DHR DOMAINS.  
 CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.  
 CC -1- SIMILARITY: BELONGS TO THE MAGUK FAMILY OF CELL JUNCTION PROTEINS.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sdb.ch/announce/>  
 CC or send an email to [license@sdb.ch](mailto:license@sdb.ch)).

Query Match 5.4%; Score 379; DB 1; Length 960;  
 Best Local Similarity 24.8%; Pred. No. 1.4e-12;  
 Matches 136; Conservative 85; Mismatches 190; Indels 138; Gaps 19;



```

Db 422 SSQSVDNHVS-----SYLGQTPASPARSPIS 450
Qy 1203 VAAQIOKLVGDRIVITCGTSTEGMTHQAVNLKNSGSIEMQVAGGVYVTHHDE 1262
Db 451 KAV-----LGDEIT-----RR 462
Qy 1263 PASSSLFTGLTSTSFQDGLGPPCKSTLREPGDLSFYGVGSPHGLPIYVKY 1322
Db 463 P-----RKVYLHNGSTGLGFNIYVG---EDGE-GITISFT 493
Qy 1323 FAKGAASEDGRKRGDQIIAIVNGQSLEGVTHEEVAAILKRTKGTVTLM 1370
Db 494 LAGGPADLSGELRKGRDRIISVNSVDLRAASHEDQAAALKNAGQAVTIV 541

RESULT 12
ID DGL1_HUMAN STANDARD: PRT; 904 AA.
AC 012959: 012958:
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE PRESYNAPTIC PROTEIN SAP97 (SYNAPSE-ASSOCIATED PROTEIN 97) (DISCS,
DE LARGE HOMOLOG 1).
GN DGL1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95024052: PubMed=7937897;
RA Lue R.A., Marfatia S.M., Branton D., Chishti A.H.;
RT "Cloning and characterization of hdlg, the human homologue of the
RT Drosophila discs large tumor suppressor binds to protein 4.1.";
RL Proc. Natl. Acad. Sci. U.S.A. 91:9818-9822(1994).
RN [2]
RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS) OF 460-555.
RX MEDLINE=96338231: PubMed=8757139;
RA Cabral J.H., Petosa C., Sutcliffe M.J., Raza S., Byron O., Poy F.,
RA Marfatia S.M., Chishti A.H., Liddington R.C.;
RT "Crystal structure of a PDZ domain.";
RL Nature 382:649-652(1996).
CC -1- FUNCTION: INTERACTS WITH THE CYTOPLASMIC TAIL OF NMDA RECEPTOR
CC SUBUNITS. ASSOCIATES WITH PROTEIN 4.1.
CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS: 1 (SHOWN HERE) AND 2; ARE
CC PRODUCED BY ALTERNATIVE SPLICING.
CC -1- SIMILARITY: CONTAINS 3 PDZ/DRH DOMAINS.
CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
CC -1- SIMILARITY: CONTAINS A GUANYLATE KINASE-LIKE DOMAIN.
CC -1- SIMILARITY: BELONGS TO THE MAGUK FAMILY OF CELL JUNCTION PROTEINS.
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: U13897; AAA50599.1; -
CC EMBL: U13896; AAA50598.1; -
CC PDB: 1PDR; 23-JUL-97.
CC MIM: 601014; -
CC InterPro: IPR000619; -
CC InterPro: IPR001452; -
CC InterPro: IPR001478; -
CC Pfam: PF00625; Guanylate_kin; 1.
CC Pfam: PF00055; PDZ; 3.
CC Pfam: PF00018; SH3; 1.
CC PROSITE: PS00856; GUANYLATE_KINASE_1; 1.
CC PROSITE: PS00502; GUANYLATE_KINASE_2; 1.
CC PROSITE: PS0106; PDZ; 3.

```

```

DR PROSITE: PS00002; SH3; 1.
KW SH3 domain; Repeat; Alternative splicing; 3D-structure.
FT DOMAIN 224 310
FT DOMAIN 319 405 PDZ 1.
FT DOMAIN 466 546 PDZ 2.
FT DOMAIN 581 651 PDZ 3.
FT DOMAIN 714 904 SH3.
FT VARSPIC 669 680 GUANYLATE KINASE.
FT ETPDDKSGSLK -> OSNRDKRKKRLPSRKKPFYKNDQS
FT EQETSDADQ (IN ISOFORM 2).
SQ SEQUENCE 904 AA; 100354 MM; B78798D6B0920D4 CRC64;

Query Match 5 4%; Score 377; DB 1; Length 904;
Best Local Similarity 29.9%; Pred. No. 1,6e-12;
Matches 119; Conservative 58; Mismatches 129; Indels 92; Gaps 14;

Qy 1002 QTPQRYRLTYRDEAPYKEEVCDTLTIEOKKPKGKLSYIGKR-----NDTGCVFS 1055
Db 208 EPTTYVNGT-----DADYEVEI-----TLE---RNGSLGFSINGTDNPHIGDSSIFIT 256
Qy 1056 DIVKGGIADPDGRLIGDQIILVNGEDVRNASQEAVALIKSLGTVTLEVRIRKAPPH 1115
Db 257 KITGGAAADGRLRYNDCTIQAENEVDVDRVTSKAVEALKKSGSYRLVYKR----- 309
Qy 1116 SERPSQTSQVSEGSLSFTFPLSGSSSTSSLESSKKNALAEIOGLRTVEKKKPTDS 1175
Db 310 --RKP-----VSE-----KMEIKLIKGP-KG 328
Qy 1176 LGISAGVGSF--LGDVPFIAMHPTGVAQOTQRLRVGDRIVITCGTSTEGMTHQAV 1233
Db 329 LGFSIAGVGNGNHPDGNSTYVTKIEGGAHKGDKLQIGDKLAVNNVCLEEVTHEEAV 368
Qy 1234 NLKNSGSIEMQVYAGGVSV-----VTGHQDEPAS---SSLSFTGLT----- 1274
Db 389 TALKNTSDFYLYKVARPTSMYNDGYAPPDITNSSQPDVNHVSPSFLGQTPASPARYS 448
Qy 1275 --STSIQDGLGPPCKSTLREPGDLSFYGVGSPHGLPIYVKYFAKGAASEDG 1332
Db 449 PVSKAVALGDEITREPRKRVYLHNGSTGLGFNIYVG---EDGE-GITISFTLAGGPADLSG 504
Qy 1333 RLKRGDQIIAIVNGQSLEGVTHEEVAAILKRTKGTVTLM 1370
Db 505 ELRKGRDRIISVNSVDLRAASHEDQAAALKNAGQAVTIV 542

RESULT 13
ID IL16_HUMAN STANDARD: PRT; 631 AA.
AC 014005: 016435; 090P18;
DT 01-NOV-1997 (Rel. 35, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE INTERLEUKIN-16 PRECURSOR (IL-16) (LYMPHOCYTE CHEMOTACTIC FACTOR)
DE (LCP).
GN IL16.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=peripheral blood;
RA Kornfeld H.;
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99145586: PubMed=9990060;
RA Bannert N., Avots A., Baler M., Serfling E., Kurth R.;
RT "GA-binding protein factors, in concert with the coactivator CREB
RT binding protein/p300, control the induction of the interleukin 16
RT promoter in T lymphocytes.";
RL Proc. Natl. Acad. Sci. U.S.A. 96:1541-1546(1999).
RN [3]

```

RP SEQUENCE OF 241-631 FROM N.A.  
 RA MEDLINE=96211141; PubMed=860984;  
 RA Bannert N., Baier M., Werner A., Kurth R.;  
 RT "Interleukin-16 or not?";  
 RL Nature 381:30-30(1996).  
 RN [4]  
 RP SEQUENCE OF 502-631 FROM N.A.  
 RC TISSUE=Periphereal Blood;  
 RA MEDLINE=94255480; PubMed=7910967;  
 RA Crilksbank W.M., Center D.M., Nisar N., Wu M., Natke B.C.,  
 RA Theodore A.C., Kornfeld H.;  
 RT "Molecular and functional analysis of a lymphocyte chemoattractant  
 factor: association of biologic function with CD4 expression.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 91:5109-5113(1994).  
 RN [5]  
 RP SEQUENCE OF 502-631 FROM N.A.  
 RA Du Y., Du G.X., Hou L.H., Wang H.;  
 RT "cDNA sequence of Interleukin-16 cloned from human PBs.";  
 RL Submitted (MAR-1998) to the EMBL/Genbank/DBJ databases.  
 CC -1- FUNCTION: STIMULATES A MIGRATORY RESPONSE IN CD4+ LYMPHOCYTES,  
 CC MONOCYTES, AND EOSINOPHILS. ALSO INDUCES T-LYMPHOCYTE EXPRESSION  
 CC OF INTERLEUKIN 2 RECEPTOR. LIGAND FOR CD4.  
 CC -1- SUBUNIT: HOMOTETRAMER.  
 CC -1- SUBCELLULAR LOCATION: SECRETED.  
 CC -1- SIMILARITY: CONTAINS 2 PDZ/DHR DOMAINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: M90391; AAD04636.1; -  
 DR EMBL: AF077011; AAD15990.1; -  
 DR EMBL: S81601; AAB36371.1; -  
 DR EMBL: AF053412; AAC12732.1; -  
 DR MIM: 603035; -  
 DR HSSP: Q12959; 1PDR.  
 DR InterPro: IPR001478; -  
 DR Pfam: PF00595; PDZ. 1.  
 DR PROSITE: PS50106; PDZ. 2.  
 KM Cytokine; Chemotaxis.  
 FT PROPEP 1.  
 FT CHAIN 1.  
 FT DOMAIN 411 496 INTERLEUKIN-16.  
 FT DOMAIN 533 618 PDZ 1.  
 FT DOMAIN 104 104 PDZ 2.  
 FT CONFLICT 233 233 E -> D (IN REF. 2).  
 FT CONFLICT 241 243 F -> L (IN REF. 2).  
 FT CONFLICT 319 319 LRL -> PRE (IN REF. 3).  
 FT CONFLICT 319 319 E -> A (IN REF. 1).  
 SQ SEQUENCE 631 AA; 66694 MW; 2A9CCED70CC8D4AC CRC64;

Query Match 4.08; Score 283; DB 1; Length 631;  
 Best Local Similarity 22.0%; Pred. No. 8.5e-08;  
 Matches 171; Conservative 106; Mismatches 306; Indels 194; Gaps 31;

624 SODVKEDEFGSMKINERYGTGLGHELEKSGSLGSLAGNDRSMSTFYGI 683  
 4 SFDTPAEPWRIISDCTINLSPINSEHN-----GHMPLQPNMSEEGTGC----- 51  
 684 DPNAGKGRGLQIADELLEINGOLYGRSHQNA--SSTIKCAPS-----KVKIIFIR 734  
 52 HPDGGPR-----LDTANGT---PKVYKASDSTVTKGPPAPAPKAPMFRQSLKGLR 99  
 735 NKDAVNAQVAPGNAVEPLPSNSENLOKKEPEPTYTSDAAVDSSEKNNVHLELPKDG 794  
 100 NRASEPRLPAPALSTGPAPASREHL--GSHIRASSSSSSIRORISSFETFGSSQLP--DKG 157  
 795 GLGAIASEDTLGSYITSLTEHGAATDGRLLKVGDOGLAVDELVGVPIEKFSILKT 854

DB 158 AGRSLQPS---SGEAPNPLGNH-----EEGR-----FSGLL-- 186  
 OY AKMTVKLTTHAENPDPSQAVPSAAGASEKKNSSOLMYPGSGPESEIRTSRSTPA 914  
 DB 187 -----GRGNA-----PLTVPO-----QPEV--LSSGSPRA 210  
 OY IFASDPATCPPIPGCETTEISKRTGLGLSTVSGSDTLGLGFIHEVEEGACKDGR 974  
 DB 211 SEARDPGVSESP-----PGRPQNKTPPPDPDLLRLSTQAEBSQRP----- 254  
 OY 975 WAGDDILEVNGIDLKATHEAINTLRQTPQVRLTVRDEAPYKEEVCOTLTLELOK 1034  
 DB 255 -----VLKMPQAR-----SEPLRRSSCETKL--LDEK 282  
 OY 1035 PCKGLGLSTVGRNDTVGVSDIVKGGIADPGRLQDQILLVNGEDVRNASQBAVAL 1094  
 DB 283 TSKLXIS-----SOVSSAVMSLCLPSSISCAQTCPCKPGASPTSSNDSMA-- 333  
 OY 1095 LKCSLGYTVLEVGRKAGPFH--SERP-----SQTQVSEGLSSTFTPLSGSTSES 1146  
 DB 334 -NGSAETSLDTG-----FSNLSELREYTGLEAKRDDDDGHSILO--SGQSVISL 383  
 OY 1147 LESSKK-----NALASEIOLGRVEMKKGPTDSLGISINGVSGPLGDVPPIFIA 1196  
 DB 384 LSSEELKILIEKVLDKATLKQIDGIIHTILHKEGAGLGSFLAG--GADLENVITYH 441  
 OY 1197 MHPFVGAQOTOKLKVGRIVITCGSTEGMHTQAVNLKNASGSIEMQVAGDVSIV 1256  
 DB 442 RVFPGNLASQESTIOKNEVLSINKSLKGTTHHDLALILKRAEPPRAVITRLTPEA 501  
 OY 1257 TGHQEPASSLSFTGLTSTSTIFQDILGPPCKSTTLERPGDLGFSIVGSGPHGLP 1316  
 DB 502 MEDLNSSTDSASASAAADVSIV--ESTAEATVCTVLEKMSGLGFSLEGKSGSLHGRK 559  
 OY 1317 IYKTVFAKGAASEDGR-LKRGDQIIANNGSLSEVTHHEANAILKR-TKGVITLAV 1371  
 DB 560 LTRIRF-KGAASEQSETVQPEDTLQIGTAMQGLTFEAMNITKALPDGVTIVT 615

RESULT 14  
 ZOL\_HUMAN STANDARD: PRT: 1736 AA.  
 AC 007157;  
 DT 01-OCT-1994 (Rel. 30, Created)  
 DT 01-OCT-1994 (Rel. 30, Last sequence update)  
 DT 01-OCT-2000 (Rel. 40, Last annotation update)  
 DE TIGHT JUNCTION PROTEIN ZO-1 (ZONA OCCUDENS 1 PROTEIN) (TIGHT JUNCTION  
 DE PROTEIN 1).  
 GN TUP1 OR ZO1.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Liver;  
 RA MEDLINE=93361541; PubMed=8395056;  
 RA Willott E., Balda M.S., Fanning A.S., Jameson B., van Itallie C.,  
 RA Anderson J.M.;  
 RT "The tight junction protein ZO-1 is homologous to the Drosophila  
 RT discs-large tumor suppressor protein of septate junctions.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 90:7834-7838(1993).  
 CC -1- FUNCTION: THE N-TERMINAL MAY BE INVOLVED IN TRANSDUCING A SIGNAL  
 CC REQUIRED FOR TIGHT JUNCTION ASSEMBLY, WHILE THE C-TERMINAL MAY  
 CC HAVE SPECIFIC PROPERTIES OF TIGHT JUNCTIONS. THE ALPHA DOMAIN  
 CC MIGHT BE INVOLVED IN STABILIZING JUNCTIONS.  
 CC -1- SUBUNIT: INTERACTS WITH OCCUDLIN, CLAUDINS AND ZO-3.  
 CC -1- SUBCELLULAR LOCATION: PERIPHERAL MEMBRANE. CYTOPLASMIC SIDE.  
 CC MOVEMENT OF ZO-1 FROM THE CYTOPLASM TO MEMBRANE IS AN EARLY EVENT  
 CC OCCURRING CONCURRENTLY WITH CELL-CELL CONTACT.  
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG FORM (SHOWN HERE) AND A  
 CC SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.  
 CC -1- TISSUE SPECIFICITY: THE ALPHA-CONTAINING ISOFORM IS FOUND IN MOST

| Query Match   | Best Local Similarity                   | Score        | DB 1 | Length |
|---|---|--------------|------|--------|
| Matches 100; conservative                                       | 68; mismatches 140; indels 182; gaps 15 | 3.5%; 20.0%; |      |        |
| 1019 KEEVCDTLTLELRKPGKGLSLVGRKN-----DTGVESDVIYKGIADPDGRLIQ      | 1071                                    |              |      |        |
| 2 EETAIWEQHTLTLHAPGFGFIAISGADNDPHFGSGETSIYSDVLKGPAA--EGOLGE     | 59                                      |              |      |        |
| 1072 GQOILLVNGEDVRRNASQF-AVAALLKCSLGT-----VLEVGRIAGFPFHSRRS     | 1121                                    |              |      |        |
| 60 NDRVAVMGVSNADNVEHAFAYOOLRRKSGKNATITRRKKKVOIIPSRPDEPVSNDNEES  | 119                                     |              |      |        |
| 1122 QTSVSEGLSFTFPLSG-----STSSLESSSKNNLASE--IQ                  | 1161                                    |              |      |        |
| 120 YDEELIHD-----PRSGRSGVNNRKSRTKIMPRDRSASRESLSPSRDRSVASQSPAK   | 171                                     |              |      |        |
| 1162 GLRTVEKKKPPDLSIGISVAGVSGSELPVFIFIAMHPTGVAAGOTKLVRGDRIVTIGG | 1221                                    |              |      |        |
| 172 PKVVLVRSKRKNEEYGLRLAS-----HIFVKRISODSLAARDGNIQEGDVLKING     | 222                                     |              |      |        |
| 1222 TSTEGTHTQAVNLKKNASGSIEMOVVAG-----GDVSV                     | 1255                                    |              |      |        |
| 223 TVTEWMSLTDAKTLTIERSGKLKMYVQDERATLLNVPLSDSIHSANASERDIDSEIOS  | 282                                     |              |      |        |
| 1256 VYGHH-----QEPASSSL-----SFTGLTST                            | 1276                                    |              |      |        |
| 283 LASDSHGRNDPRPRRSRSRSPDQRSEPSDHSRHSPQDSNGSLSRSDERISKPAVST    | 342                                     |              |      |        |

```

Oy 1277 SI-PODDLGPPQCKSITLERGP-----1297
Db 343 PVKADHDPTVEVETVETERNKEKOTPSLPERKPVYAQVGNOMIYLSVHLAVSYLIOLMK 402
Oy 1298 -----DELGSIVGGYGSPPHDLPIYKVTYFAKGAASEDRLRKGDII 1341
Db 403 MGLFASPKKLVKFRKGDVGLRLAG-----NDVGIFVAGVLEDSPPAKRG-LEGGDIL 456
Oy 1342 AVNGOSLEGVTHREAVAIL 1360
Db 457 RYNNVDFTNIIREEVALEL 475

RESULT 15
Z03_HUMAN
ID Z03_HUMAN STANDARD; PRT; 933 AA.
AC 095049;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE TIGHT JUNCTION PROTEIN ZO-3 (ZONA OCCUDENS 3 PROTEIN) (TIGHT JUNCTION
DE PROTEIN 3).
GN TPJ3 OR ZO3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
[1]
SEQUENCE FROM N.A.
R lamerdin J.E., McCreedy P.M., Skowronski E., Viswanathan V.,
R Burkhardt-Schultz K., Brower A., Gordon L., Dias J., Ramirez M.,
R Stillwagen S., Phan H., Velasco N., Do L., Regala M., Terry A.,
R Ganes J., Dangsan L., Erlar A., Christensen M., Georgescu A.,
R Avila J., Liu S., Attix C., Andreise T., Frankheim M.,
R Amico-Keller G., Coerfield J., Duarte S., Lucas S., Bruce R.,
R Thomas P., Quan G., Kronmiller B., Arellano A., Sanders C., Ow D.,
R Nolan M., Tjong S., Kobayashi A., Olsen A.S., Carraro A.V.;
RA "Sequence analysis of a 3.5 kb contig in human 19p13.3 containing a
RT serine protease gene cluster";
RL Submitted (NOV-1998) to the EMBL/Genbank/DBJ databases.
CC -1- SUBUNIT: INTERACTS WITH OCCUDLIN, CLAUDINS AND ZO-1 (BY
CC -1- SIMILARITY).
CC -1- SIMILARITY: CONTAINS 3 PDZ/DHR DOMAINS.
CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
CC -1- SIMILARITY: CONTAINS A GUANYLATE KINASE-LIKE DOMAIN.
CC -1- SIMILARITY: BELONGS TO THE MAGUK FAMILY OF CELL JUNCTION PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC DR EMBL; AC005954; AAC72274.1; ALT_INIT.
CC DR HSSP; P31016; 1BPE.
CC DR InterPro; IPR000619; -.
CC DR InterPro; IPR001452; -.
CC DR InterPro; IPR001478; -.
CC DR Pfam; PF00595; PDZ; 3.
CC DR PROSITE; PS00625; Guanylate_kin; 1.
CC DR PROSITE; PS00856; GUANYLATE_KINASE_1; FALSE_NEG.
CC DR PROSITE; PS50052; GUANYLATE_KINASE_2; 1.
CC DR PROSITE; PS50002; SH3; 1.
CC DR PROSITE; PS50106; PDZ; 3.
CC KW Tight junction; SH3 domain; Repeat; Membrane.
CC FT DOMAIN 11 93 PDZ 1.
CC FT DOMAIN 195 272 PDZ 2.
CC FT DOMAIN 394 460 PDZ 3.
CC FT* DOMAIN 489 563 SH3.
CC FT DOMAIN 675 775 GUANYLATE_KINASE.

```

SO SEQUENCE 933 AA; 102800 MW; 680298CFD0615847 CRC64;

Query Match 3.5%; Score 242; DB 1; Length 933;

Best Local Similarity 22.8%; Pred. No. 2.3e-05; Matches 121; Conservative 62; Mismatches 155; Indels 192; Gaps 20;

```
QY 932 TLEISKG-RTGLGSLIVGSDTLGAFTHIEYEGACKDGLWAGDQILEVNGIDLK 990
| : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 11 TATLSKDPRRRGIAISGRDRPGSMVSDVPGPA--EGRLQTDGHIVANGVSMEN 68
| : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 991 ATHDEAINVLR-----OTPORVRLTLR-----DEAPYKEEYV----- 1023
| : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 69 ATSAFAIQILKTCMTMANITVKRRRIHLPATKASPSSPGRDSDDEDGQPQRVEEVDQR 128
| : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1024 -----CDYLT-----EL 1031
| : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 129 GYDDSSSGSGSRSDERRRRPRGRAGSHGRSPGGSGSEANGLALVSGFKRLPRDV 188
| : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1032 QKKPGKGLSLIVGKRD-----TGVPVSDIVKGIADPDRLIQDQIILLNGED 1082
| : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 189 QMKPYK---SVLVKRRDSEEGVNLGSIQIFIKHTDGLAARHRLQEGDLILQINGVS 244
| : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1083 VRNASQEAVALKCSLCTVLEVR-----1108
| : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 245 SONSLNDTRLRIEKSGLSLVLRDGOFLVNIIPAVSDSDSSPLEEGVTMADEMSSP 304
| : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1109 -----IKAGPFHSERRP--SOTSQV-----SEGLSFPPLSGS 1141
| : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 305 PADISDLASELSQAPPSHIIPPRHAKNSPESQSDSPVESRRLRRESSVDRTISEPDE 364
| : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1142 STSESLSSS-----KKNALASEIOG---LRTYEMKKGPTDSLGISIAGVGSPLDVP 1192
| : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 365 QRSELPRESSYDIYVPSQSMEDRGVSPDTRVYVFLKG--KSIQLRLAG-----NDVG 417
| : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1193 IFIMMHFTGYAAQTKLRVGDRTVITIGTSTEGHTQAVNLKNAGSIEMQVYAG-- 1250
| : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 418 IFVSGVQ-AGSPADGQIQEGDQILQVNDVPQNLTRREAVQFLGLPGEEMELVYQK 476
| : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1251 -----GDVSVVTGHHO--EPASSSLFTGILTSTSIFO--DGLP 1285
| : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 477 QDIFMKWQSVKRGDSFYIRTHFELEP--SPPSGLGFTRGDVFHVLDTLHP 524
| : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

Search completed: July 12, 2001, 14:43:17  
Job time: 210 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: July 12, 2001, 14:39:27 ; Search time 64.35 Seconds

(without alignments) updates/sec  
2822.918 Million cell

Title: US-09-502-698-1

Perfect score: 6990

Sequence: 1 MWCRRTPPTQSELDLDLDCIEETKPHVDLGEFISSTEDPVLMTAGOSTEEV 1373

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 425026 seqs, 132305027 residues

Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

## Database :

SPTREMBL\_16:\*

- 1: sp.archaea:\*
- 2: sp.bacteria:\*
- 3: sp.fungi:\*
- 4: sp.human:\*
- 5: sp.invertebrate:\*
- 6: sp.mammal:\*
- 7: sp.mbc:\*
- 8: sp.organelle:\*
- 9: sp.phage:\*
- 10: sp.plant:\*
- 11: sp.todent:\*
- 12: sp.unclassified:\*
- 13: sp.vertebrate:\*
- 14: sp.virus:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score  | Query Match | Length | DB ID     | Description        |
|------------|--------|-------------|--------|-----------|--------------------|
| 1          | 6735   | 96.4        | 2042   | 4 075970  | 075970 homo sapien |
| 2          | 5811   | 83.1        | 2054   | 11 055164 | 055164 rattus norv |
| 3          | 5617   | 80.4        | 2055   | 11 0921K3 | 0921K3 mus musculu |
| 4          | 2297.5 | 32.9        | 526    | 11 008783 | 008783 mus musculu |
| 5          | 2252   | 32.2        | 453    | 4 043798  | 043798 homo sapien |
| 6          | 1820.5 | 26.0        | 1134   | 4 09H3N9  | 09H3N9 homo sapien |
| 7          | 1550.5 | 22.2        | 1552   | 4 060833  | 060833 homo sapien |
| 8          | 1525.5 | 21.8        | 1582   | 4 043742  | 043742 homo sapien |
| 9          | 1469.5 | 21.0        | 612    | 11 070471 | 070471 mus musculu |
| 10         | 1390.5 | 19.9        | 1524   | 4 015249  | 015249 homo sapien |
| 11         | 1044.5 | 14.9        | 2208   | 5 009515  | 009515 caenorhabdi |
| 12         | 589    | 8.4         | 728    | 11 070263 | 070263 mus musculu |
| 13         | 564.5  | 8.1         | 665    | 5 023823  | 023823 calliphora  |
| 14         | 556    | 8.0         | 674    | 5 028BV3  | 028BV3 drosophila  |
| 15         | 554    | 7.9         | 674    | 5 024008  | 024008 drosophila  |
| 16         | 534    | 7.6         | 1012   | 5 021075  | 021075 caenorhabdi |
| 17         | 517.5  | 7.4         | 1551   | 4 014160  | 014160 homo sapien |
| 18         | 507.5  | 7.3         | 2460   | 11 064512 | 064512 mus musculu |
| 19         | 507    | 7.3         | 2484   | 6 028006  | 028006 bos taurus  |

|    |       |     |      |           |                    |
|----|-------|-----|------|-----------|--------------------|
| 20 | 453.5 | 6.5 | 1256 | 4 075085  | 075085 homo sapien |
| 21 | 443   | 6.3 | 2766 | 11 09QZB8 | 09QZB8 rattus norv |
| 22 | 441   | 6.3 | 1150 | 4 09HBC4  | 09HBC4 homo sapien |
| 23 | 441   | 6.3 | 1756 | 5 09VBE4  | 09VBE4 drosophila  |
| 24 | 432   | 6.2 | 1125 | 4 09H2V6  | 09H2V6 homo sapien |
| 25 | 431.5 | 6.2 | 1179 | 11 09JK71 | 09JK71 rattus norv |
| 26 | 429.5 | 6.1 | 1126 | 11 09EOJ9 | 09EOJ9 mus musculu |
| 27 | 426   | 6.1 | 1455 | 4 060510  | 060510 homo sapien |
| 28 | 425   | 6.1 | 874  | 4 09HCD8  | 09HCD8 homo sapien |
| 29 | 424   | 6.1 | 871  | 5 09XZ35  | 09XZ35 drosophila  |
| 30 | 423.5 | 6.1 | 1278 | 4 060434  | 060434 homo sapien |
| 31 | 422   | 6.0 | 871  | 5 09NB04  | 09NB04 drosophila  |
| 32 | 409   | 5.9 | 1112 | 11 09WV01 | 09WV01 mus musculu |
| 33 | 399.5 | 5.7 | 1131 | 5 018165  | 018165 caenorhabdi |
| 34 | 398.5 | 5.7 | 1277 | 11 088382 | 088382 rattus norv |
| 35 | 397.5 | 5.7 | 1114 | 11 09RZ71 | 09RZ71 rattus norv |
| 36 | 385   | 5.5 | 960  | 5 09VY25  | 09VY25 drosophila  |
| 37 | 385   | 5.5 | 1171 | 11 054893 | 054893 mus musculu |
| 38 | 381   | 5.5 | 950  | 11 09WU36 | 09WU36 rattus norv |
| 39 | 379.5 | 5.4 | 1043 | 11 09WTW1 | 09WTW1 rattus norv |
| 40 | 377.5 | 5.4 | 822  | 11 088961 | 088961 rattus norv |
| 41 | 376   | 5.4 | 1322 | 11 09QZP6 | 09QZP6 mus musculu |
| 42 | 370   | 5.3 | 1027 | 4 09H1X7  | 09H1X7 homo sapien |
| 43 | 368   | 5.3 | 927  | 11 062402 | 062402 mus musculu |
| 44 | 357   | 5.1 | 1112 | 11 097879 | 097879 rattus norv |
| 45 | 353   | 5.1 | 1464 | 5 096782  | 096782 drosophila  |

## ALIGNMENTS

| RESULT                     | ID   | PRELIMINARY: | PRT: | 2042 AA. |
|----------------------------|--|--------------|------|----------|
| 075970                     | 075970   | PRELIMINARY: | PRT: | 2042 AA. |
| AC                         | 075970:  |              |      |          |
| DT                         | 01-NOV-1998 (TREMBLrel. 08, Created)                               |              |      |          |
| DT                         | 01-NOV-1998 (TREMBLrel. 08, last sequence update)                  |              |      |          |
| DT                         | 01-MAR-2001 (TREMBLrel. 16, last annotation update)                |              |      |          |
| DE                         | MULTI PDZ DOMAIN PROTEIN MUPP1.                                    |              |      |          |
| GN                         | MUPP1.   |              |      |          |
| OS                         | Homo sapiens (Human).  |              |      |          |
| OC                         | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  |              |      |          |
| OC                         | Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.          |              |      |          |
| OX                         | NCBI_Taxid=9606;   |              |      |          |
| RN                         | [1]  |              |      |          |
| RP                         | SEQUENCE FROM N.A.   |              |      |          |
| RC                         | TISSUE=BRAIN;  |              |      |          |
| RA                         | Eng L., Kravitsky G., Clapham P.E.;                                |              |      |          |
| RT                         | "Human homolog of MUPP1 protein."                                  |              |      |          |
| RL                         | Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.            |              |      |          |
| DR                         | EMBL; AF093419; AAC61870.1; -.                                     |              |      |          |
| DR                         | HSSP; Q12959; 1PDR.  |              |      |          |
| DR                         | InterPro; IPR001478; -.  |              |      |          |
| DR                         | pfam; PF00595; PDZ; 13.  |              |      |          |
| DR                         | SMART; SM00228; PDZ; 1.  |              |      |          |
| SQ                         | SEQUENCE 2042 AA; 218595 MW; F0E1DD029749A762 CRC64;               |              |      |          |
| Query Match                | 96.4%; Score 6735; DB 4; Length 2042;                              |              |      |          |
| Best local Similarity      | 92.6%; Pred. No. 0;  |              |      |          |
| Matches 1336; Conservative | 6; Mismatches 3; Indels 98; Gaps 2;                                |              |      |          |
| QY                         | 1 MWCRRTPPTQSELDLDLDCIEETKPHVDLGEFISSTEDPVLMTAGOSTEEV 60           |              |      |          |
| DB                         | 628 MWCRRTPPTQSELDLDLDCIEETKPHVDLGEFISSTEDPVLMTAGOSTEEV 687        |              |      |          |
| QY                         | 61 QAPLAMEAGIHIETKSGKSGFSLDYOPIDPASAIVIIIRSYPGIGIAEKORLLP 120      |              |      |          |
| DB                         | 688 QAPLAMEAGIHIETKSGKSGFSLDYOPIDPASAIVIIIRSYPGIGIAEKORLLP 747     |              |      |          |
| QY                         | 121 GRLMFVNDVNLNENSSLEAEVALKAPSGTVRIQVAKPLPLSPREGYVSAKEDSFLYPP 180 |              |      |          |
| DB                         | 748 GRLMFVNDVNLNENSSLEAEVALKAPSGTVRIQVAKPLPLSPREGYVSAKEDSFLYPP 807 |              |      |          |

```

QY 181 HSCBAGIADKPLRADLALVTNDADLVDESTESPSPENDSIYTOASTLSLHGSSC 240
   |||||
Db 808 HSCBAGIADKPLRADLALVTNDADLVDESTESPSPENDSIYTOASTLSLHGSSC 867
QY 241 GDGLNTGSSLPSPKDVYIENSCDPLDLHMSLEELYTQNLERODENTPSVDISMGPAS 300
   |||||
Db 868 GDGLNTGSSLPSPKDVYIENSCDPLDLHMSLEELYTQNLERODENTPSVDISMGPAS 927
QY 301 GFTINDTPANAIOOYECENTIYWTESHLPSEYISSAELPSVLPDSAGKSEHLLEOSS 360
   |||||
Db 928 GFTINDTPANAIOOYECENTIYWTESHLPSEYISSAELPSVLPDSAGKSEHLLEOSS 987
QY 361 LACNAECMLONVKESERTINIAKGNSSIGMTVSNKDGIMIVSIIIGCAISRDR 420
   |||||
Db 988 LACNAECMLONVKESERTINIAKGNSSIGMTVSNKDGIMIVSIIIGCAISRDR 1047
QY 421 IAIIDCJLISNEESTISVTNAQARAMLRRHSLIGPDIKITVPAHLEEFKISLGQSGR 480
   |||||
Db 1048 IAIIDCJLISNEESTISVTNAQARAMLRRHSLIGPDIKITVPAHLEEFKISLGQSGR 1107
QY 481 VMAIDFSSYGRDIPPEBEGESELONTATSNMNPRAVELMREPSKSLGISTV 540
   |||||
Db 1108 VMAIDFSSYGRDIPPEBEGESELONTATSNMNPRAVELMREPSKSLGISTV 1167
QY 541 GGRMGSRLSNGEVMRGIFIKHVLSDPAGKNGTLKPDRIVE----- 583
   |||||
Db 1168 GGRMGSRLSNGEVMRGIFIKHVLSDPAGKNGTLKPDRIVEYDGMRLDASHGQAVEA 1227
QY 584 -----APSGSES 590
   |||||
Db 1228 IRKAGNPVFNWOSIINRPRKSPLELHNLPRYNFSSTNPADSLQINADKAPSOSES 1287
QY 591 EPEKAPLCSVPPPPSAFAEMGSDHTOSASKISQDYKEDDFGYSMKINERKGTLTGE 650
   |||||
Db 1288 EPEKAPLCSVPPPPSAFAEMGSDHTOSASKISQDYKEDDFGYSMKINERKGTLTGE 1347
QY 651 LHMTELEKSHGSLGSLAGNKRSMVSFYIGIDPNGAAGDGRLOLADLELEINGOITLY 710
   |||||
Db 1348 LHMTELEKSHGSLGSLAGNKRSMVSFYIGIDPNGAAGDGRLOLADLELEINGOITLY 1407
QY 711 GRSQONASSIICAPSKYKIIIFIRNKDAVNOMAVCPGNAVEPLPSNSFNONKTEPEVPT 770
   |||||
Db 1408 GRSQONASSIICAPSKYKIIIFIRNKDAVNOMAVCPGNAVEPLPSNSFNONKTEPEVPT 1467
QY 771 TSDAAYVLSSEKKNYQHLLEPKDQGLGIAISEEDTLGVIISKSLTEHGVATDRLKYGD 830
   |||||
Db 1468 TSDAAYVLSSEKKNYQHLLEPKDQGLGIAISEEDTLGVIISKSLTEHGVATDRLKYGD 1527
QY 831 QILAVDEIVGVPIEKFISLTKAKMTVKLTIAENPDQAVPSAAGASGERKKNSSQS 890
   |||||
Db 1528 QILAVDEIVGVPIEKFISLTKAKMTVKLTIAENPDQAVPSAAGASGERKKNSSQS 1587
QY 891 LAMVQSSPESERTIRNRSSTPAIFASDPATCPIPCCEETIISGRGLGISTYGS 950
   |||||
Db 1588 LAMVQSSPESERTIRNRSSTPAIFASDPATCPIPCCEETIISGRGLGISTYGS 1647
QY 951 DTLGAFIIEHVEYEGAAKCDGRIMAGDQILEVNGIDLKRAKTHDEAIVLROTQORVLT 1010
   |||||
Db 1648 DTLGAFIIEHVEYEGAAKCDGRIMAGDQILEVNGIDLKRAKTHDEAIVLROTQORVLT 1707
QY 1011 LYRDEAPYKEEEVCDTLTIELQKRPKGGLGISTYGRNDYGVFSDIYKGGIADPGRRLI 1070
   |||||
Db 1708 LYRDEAPYKEEEVCDTLTIELQKRPKGGLGISTYGRNDYGVFSDIYKGGIADPGRRLI 1767
QY 1071 QGDDIILVNGEDVANAQOENAAALLKCSIGTYLLEVGRIRAGPHERRPSQTSQVSEGS 1130
   |||||
Db 1768 QGDDIILVNGEDVANAQOENAAALLKCSIGTYLLEVGRIRAGPHERRPSQTSQVSEGS 1808
QY 1131 LASFTEPLSGSSTSESSSKKNALASEIOGLTVEAKKGPDTSLGISTAGVGSPLGD 1190
   |||||
Db 1809 -----GSSTSESSSKKNALASEIOGLTVEAKKGPDTSLGISTAGVGSPLGD 1859

```

```

QY 1191 VPIFIAMHPTGVAAGTOKLRVGDRIYICSTEGMHTQAVMLNKASSIEMQVAV 1250
   |||||
Db 1860 VPIFIAMHPTGVAAGTOKLRVGDRIYICSTEGMHTQAVMLNKASSIEMQVAV 1919
QY 1251 GDVSVVTGHHOEPASSSLSTFTGLTSTIFODDLCPPOCKSTTLERPGDGLFSTVGYGS 1310
   |||||
Db 1920 GDVSVVTGHHOEPASSSLSTFTGLTSTIFODDLCPPOCKSTTLERPGDGLFSTVGYGS 1979
QY 1311 PHGDLPIYKTVFAKGAASEGRLKRGDQITAVNGOSLEGVTHEEAVAILKRTKGTVTLM 1370
   |||||
Db 1980 PHGDLPIYKTVFAKGAASEGRLKRGDQITAVNGOSLEGVTHEEAVAILKRTKGTVTLM 2039
QY 1371 VLS 1373
   |||
Db 2040 VLS 2042

RESULT 2
ID 05164 PRELIMINARY; PRT; 2054 AA.
AC 05164;
DT 01-JUN-1998 (Tremblrel. 06, Created)
DT 01-JUN-1998 (Tremblrel. 06, Last sequence update)
DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
DE MULTI PDZ DOMAIN PROTEIN 1.
GN MUPPL.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sclurognathu; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RX MEDLINE=9819685; PubMed=9537516;
RA Ulmer C., Schumck K., Fluge A., Lubbert H.;
RT "Cloning and characterization of MUPPL, a novel PDZ domain protein.";
RL FBS Letc. 424:63-68(1998).
DR EMBL; AJ001320; CA04681.1; -.
DR HSSP; Q12959; 1PDR.
DR InterPro; IPR001478; -.
DR Pfam; PF00595; PDZ; 13.
DR SMART; SM00228; PDZ; 1.
SQ SEQUENCE 2054 AA; 218590 MW; 44BD3F428801F78F CRC64;

Query Match 83.1%; Score 5811; DB 11; Length 2054;
Best Local Similarity 79.5%; Pred. No. 0;
Matches 1147; Conservative 102; Mismatches 116; Indels 78; Gaps 5;

QY 1 MVCCRTVPTTQSELDLDCDIELEKPHVDLGEFISSEEDBPVLAAMDAGQSTEEV 60
   |||||
Db 620 MVCCRTVPTTQSELDLDCDIELEKPHVDLGEFISSEEDBPVLAAMDAGQSTEEV 679
QY 61 QAPLAMEAGQIHLLEKSGKGLFSLDQDPIDPASTVIYIISLVPYGGIAEKDGLLP 120
   |||||
Db 680 QAPLAMEAGQIHLLEKSGKGLFSLDQDPIDPASTVIYIISLVPYGGIAEKDGLLP 739
QY 121 GDRLMFNDVNLNLSLEAEVALKAPSGTVRIGVAKPLPLSPSEGYVSAKEDSFYYP 180
   |||||
Db 740 GDRLMFNDVNLNLSLEAEVALKAPSGTVRIGVAKPLPLSPSEGYVSAKEDSFYYP 799
QY 181 HSCBAGIADKPLRADLALVTNDADLVDESTESPSPENDSIYTOASTLSLHGSSC 240
   |||||
Db 800 HSCBAGIADKPLRADLALVTNDADLVDESTESPSPENDSIYTOASTLSLHGSSC 859
QY 241 GDGLNTGSSLPSPKDVYIENSCDPLDLHMSLEELYTQNLERODENTPSVDISMGPAS 300
   |||||
Db 860 GDGLNTGSSLPSPKDVYIENSCDPLDLHMSLEELYTQNLERODENTPSVDISMGPAS 918
QY 301 GFTINDTPANAIOOYECENTIYWTESHLPSEYISSAELPSVLPDSAGKSEHLLEOSS 360
   |||||
Db 919 GFTINDTPANAIOOYECENTIYWTESHLPSEYISSAELPSVLPDSAGKSEHLLEOSS 974

```

|    |      |  |      |
|----|------|--|------|
| Qy | 361  | LACAEVLMONJYKSEFEETIINAAGNSLIGMTVAANDGIMTVRSTIHGALSPRCR          | 420  |
| Db | 975  | LVSDAEVTLQMSQOEFEETVTIJAAGSSSLGWTVAANDGLGVITVRSTIHGALSTRGR       | 1034 |
| Qy | 421  | IAIDCCLISTINEESTIENVTAQARALRRISLJLGPDKITTYPAHLEEFKJSLQOQSGR      | 480  |
| Db | 1305 | IAYDCCILISTINEESTIETLTAQAARALRRISLJLGPDKITTYPAHLEEFKJSLQOQSGR    | 1094 |
| Qy | 481  | VMALDFSSYTGRIPELPEREGEGESELELONTAVSNMNOBRVLEWRPSSLGISIV          | 540  |
| Db | 1095 | IMALDFSSYTGRIPELPEREGEGESELELONTAVSNMNOBRVLEWRPSSLGISIV          | 1154 |
| Qy | 541  | GGRRMGSRLSNGEVMKICIFKHYLEDSPPAKNGTLPGRIVYE-----                  | 583  |
| Db | 1155 | GGRRMGSRLSNGEVMKICIFKHYLEDSPPAKNGTLPGRIVYEVMGMDLRASHOAVEA        | 1214 |
| Qy | 584  | -----APSOSES 590   |      |
| Db | 1215 | IRKAGSVYVMEVOSIYVRPKSPLPSLPHSLYPKCSFSTNPPASLOLTSUKAPSOSES        | 1274 |
| Qy | 591  | EPEKAFLCSVPPPPAFAMGSDHTQSSASKISQDVNDEDEFEFGSMKNIRERYTTLGE        | 650  |
| Db | 1275 | ESEKATLCSPVSSSPSVFSSMSDQAQVSAITYVAEDKEDKEDFGYSWKNIQERYTTLGQ      | 1334 |
| Qy | 651  | LHMTLEKSHSGJGLSLAGKKDRSRMSVFIVGIDPNGAAGKQGRUQIADLELITNGOILY      | 710  |
| Db | 1335 | LHMTLEKSHSGJGLSLAGKKDRSRMSVFIVGIDPNGAAGRGRUQIADLELITNGOILY       | 1394 |
| Qy | 711  | GRSHQNNASITIKCAPSVKIIIFIRNDAVNOAMVCGNAVPELPSNENLOKMEPEVYT        | 770  |
| Db | 1395 | GRSHQNNASITIKCAPSVKIIIFIRNDAVNOAMVCGSADAPLPTSESPOKMEPEVYT        | 1454 |
| Qy | 771  | TSDAANDVLSFKKNVQHLLEPKDQGGJGIAISEEDTLISGYIKSLTSEHGYAATDGLKXGD    | 830  |
| Db | 1455 | TSAAVADLSTLTNYHLELKPQDGGJGIAICEEDTLINGYIKSLTERGGAANDGLKXGD       | 1514 |
| Qy | 831  | QILAANDEIVYGRPIEKFITSLTKAATKYKLTTHAENPDQAQVPSAAGASGEKKNSSQS      | 890  |
| Db | 1515 | RILAVDELVAGCPRIEKFITSLTKAATKYKLTGVAENPDQAQVPSAAVATSEKKNSSQS      | 1574 |
| Qy | 891  | LMVQSSPPEPESIRNRSRSPATRASDPANCPIIRGCEETIETLSKGTGIGLSIVGS         | 950  |
| Db | 1575 | PAPV---APDELPETPSRKSSTPATRASDPANCPIIRGCEETIETLSKGTGIGLSIVGS      | 1631 |
| Qy | 951  | DTLLGAETIHEVYEEGAACDKDGRILMAGDQILEVNGIDLKAKTHADAIVNLRQTPQVRVLT   | 1010 |
| Db | 1632 | DTLLGAIITIIHEVYEEGAACDKDGRILMAGDQILEVNGIDLKAKTHADAIVNLRQTPQVRVLT | 1691 |
| Qy | 1011 | LYRDEAPYKEEBCDITLTELQKPKGKGLSIVGKRNDTGVFVSDIYKGIADPDGRILT        | 1070 |
| Db | 1692 | LYRDEAPYKEEBCDITLTELQKRPKGKGLSIVGKRNDTGVFVSDIYKGIADPDGRILM       | 1751 |
| Qy | 1071 | QGGDILLVNGEDVNAQOEVAALLKCSIGTLYLENGRIKAPRPISESRPQSTQSOVSGS       | 1130 |
| Db | 1752 | QGGDILLVNGEDVNAQOEVAALLKCSIGTLYLENGRIKAPRPISESRPQSTQSOVSGS       | 1811 |
| Qy | 1131 | LSSETFLPSGSSSTSESSLESSSKKNALASEIOLGRTYEMKRCPTDLSIGISIAGVSPGLD    | 1190 |
| Db | 1812 | LSSETFLPSGSSSTSESSLESSSKKNALASEIOLGRTYEMKRCPTDLSIGISIAGVSPGLD    | 1871 |
| Qy | 1191 | VPIETAMNHPTGVAOQOKRQVDRIVYTICGISTESGTHPTQAVNLKKNASGSIEMQVYAG     | 1250 |
| Db | 1872 | VPIETAMNHPTGVAOQOKRQVDRIVYTICGISTESGTHPTQAVNLKKNASGSIEMQVYAG     | 1931 |
| Qy | 1251 | GDVSVWGHNOEPRSSSLSTGJLSTISJRODDIGRPOCKSTILBNGPGLSGSYVGGVS        | 1310 |
| Db | 1932 | GDVSVWGHNOEPRSSSLSTGJLSTISJRODDIGRPOCKSTILBNGPGLSGSYVGGVS        | 1991 |
| Qy | 1311 | PHGGLPIYVTVFAKGAASEDGRILKRGDOITIANQOSELGCVHNEEVAIVLAKTRKTGYIM    | 1370 |
| Db | 1992 | PHGGLPIYVTVFAKGAASEDGRILKRGDOITIANQOSELGCVHNEEVAIVLAKTRKTGYIM    | 2051 |
| Qy | 1371 | VLS 1373   |      |

```

DB      2052 VLS 2054
RESULT      3
09Z1K3
ID      09Z1K3      PRELIMINARY;      PRT: 2055 AA.
AC      09Z1K3;
DT      01-MAY-1999 (TReMBLrel, 10, Created)
DT      01-MAY-1999 (TReMBLrel, 10, Last sequence update)
DT      01-MAR-2001 (TReMBLrel, 16, Last annotation update)
DE      MULTIPLE PDZ DOMAIN PROTEIN.
        MPDZ.
GN      Mus musculus (Mouse).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
        Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
OX      NCBI_TaxID=10090;
        [1]
RN      SEQUENCE FROM N.A.
RP      STRAIN-C57/BL6 X CBA F1; TISSUE-WHOLE BRAIN;
RA      Simpson E.H., Suffolk R., Jackson I.J.;
RT      Identification and mapping of mouse Multiple PDZ domain protein,
        Modz".
RL      Submitted (JAN-1999) to the EMBL/GenBank/DBDJ databases.
DR      EMBL; AJ131869; CAA10523.1; -.
DR      HSSP; Q12959; 1PDR.
DR      MGD; MGI:1343489; Mpdz.
DR      InterPro: IPR001478; -.
DR      Pfam; PF00595; PDZ, 13.
DR      SMART; SM00228; PDZ; 1.
SO      SEQUENCE      2055 AA; 218968 MW; E1A38EE634CC20E4 CRC64;

```

| Query Match           | 80.4%            | Score 5617  | DB 11     | Length 2055 |
|-----------------------|------------------|---|-----------|-------------|
| Best Local Similarity | 77.1%            | Pred. No. 0   |           |             |
| Matches 1112          | Conservative 114 | Mismatches 139  | Indels 78 | Gaps 5      |
| QY                    | 1                | MVCCRTVPPTTQSEISLDCIDIELTEKPHNVLGFISSSEFEDPVLAMTADAGSTEEV     | 60        |             |
| DB                    | 621              | MVCCRTVPPIALISMDSDSLINDLETERPHIDLGEFISSETEFEDPMLMSDVQNAEEL    | 680       |             |
| QY                    | 61               | QAPLAMEAGIOHIELEKSGKLGFSILDYODPIDPASTVYIIIRSLVPGGIAEKDGRLLP   | 120       |             |
| DB                    | 681              | QPLAMEWAGGOSIELEKSGRGLGFSIIDYODPIDPANTVIYIRSLVPGGIAEKDGRLLP   | 740       |             |
| QY                    | 121              | GRLLPFVNDIENLSLEAEVALGAPSGCTRIGVAKRLPLSPGGYSAKDSRLTYRP        | 180       |             |
| DB                    | 741              | GRLLPFVNDIENLSLEAEVALGAPSGMVRIGAKPLPSPEGYVSAEDAFRLCS          | 800       |             |
| QY                    | 181              | HSCEAGIADLDPFLFRDLALVGTNADVLDESSTFESYPENDSIYSTASILSLGSSC      | 240       |             |
| DB                    | 801              | HACKSGSLDKRLFRADLALIDTPRAESIASESFQSPENDSVYSTQASIFSLMDGTC      | 860       |             |
| QY                    | 241              | GDGLIAGSSLPSSPPKPDVYENSCDPYLDLHMSLEELYTONLLERODEMTPSDISMGPPAS | 300       |             |
| DB                    | 861              | SGGMNMGPSLSSPPKPDYVSSS-EVYLGLHLSTLEELYTONLLORHAGSSSDMKPTPTF   | 919       |             |
| QY                    | 301              | GPTINDYPPANALIEOYECEENTIVMTSHLSPEVSIASLPLVPLPDSACKSGSHLEQSS   | 360       |             |
| DB                    | 920              | GXISIDYTTTNGEOXYQOANVPANVHSDLPSSN-IJSTFELASL--AAVAPRYLTDQTF   | 975       |             |
| QY                    | 361              | IACNAECVLAQWVSKSEFPRTINIAKGNSSLGMTVYSANKDGLMIVRSIIHGGAISRDR   | 420       |             |
| DB                    | 976              | LASDAESVYLOGMSQEAFFRPYTIAKGSSSLGMTVYSANKDGLVYIRSLIIHGGAISRDR  | 1035      |             |
| QY                    | 421              | IAIGDCIISINEESTISVTNAOARALRRHSLIGPIDKITVYPAPHLLEEFKISLGOOSCR  | 480       |             |
| DB                    | 1036             | IAYGDCIISINEESTISLTNAOARALRRHSLIGPIDKITVYPAPHLLEEFVRSFGQOAGG  | 1095      |             |
| QY                    | 481              | WALDIFSSYTRDRIPELPEREGEGEESIELONTAASNNQRRRVLAKEPSSKISLSTIV    | 540       |             |
| DB                    | 1096             | WALDIFSSYTRDRIPELPEREGEGEESIELONAAASNNQRRRVLAKEPSSKISLSTIV    | 1155      |             |

|        |   |  |      |
|--------|---|--|------|
| QY     | 541   | GGGCGSGSLTSGEYMRGIFIKHYLEDSPPAKNTLKGRIYE-----                  | 583  |
| Db     | 1136  | GGCGSGSLTSGEYMRGIFIKHYLEDSPPAKNTLKGRIYE-----                   | 1215 |
| QY     | 584   | -----ADSGSES 590   |      |
| Db     | 1216  | IRKAGNPVYEWVQSIIRPKRSP,LSPLSHLYPKYSFSSTNPFADSLQTLTDOARSQSES    | 1275 |
| QY     | 591   | EPEKAPLCSTVPPPPSAFAEMGSDPHQSSAKRISODVDEDEFEYGMKNIIRRYGTGLGE    | 650  |
| Db     | 1276  | ETEKPALCNVPPSSPVSFSEMGSDCAQPSATVASEDEDEDEFGSMKNIDGRYSGLTGQ     | 1335 |
| QY     | 651   | LHMLELEKHSGLGLSLAGKKDRSRMSVFIVGIDPNGAAGKDGRLQIADLELLEINGQLTY   | 710  |
| Db     | 1336  | LHVIIELEGGOSGLGLSLAGKKDRTRMSVFIVGIDPNGAAGRGRLQIADLELLEINGQLTY  | 1395 |
| QY     | 711   | GRSHONASSIIKCAPSVKTIIFTRNDVAVQMAVCGNANVEDLPNSNEQLQKNEPIVY      | 770  |
| Db     | 1336  | GRSHONASSIIKCAPSVKTIIFTRNDVAVQMAVCGNANVEDLPNSNEQLQKNEPIVY      | 1455 |
| QY     | 771   | TSDAAYDLSSFKYQNHLEPRDGGGLGIALSEEDTSGYIILSTLEHGAATDGLRKG        | 830  |
| Db     | 1456  | TSASAADSLTDVYQLELPRDGGGLGIALCEEDTINGVIESLETHEGGAADGLRKG        | 1515 |
| QY     | 831   | QILAVDEIYVGYIEKFIISLTKAKTKYKLTTHAENPDSQAVPSAAGAASEKKNSSQS      | 890  |
| Db     | 1516  | HLAVDEYVAGCVERFISLTKATAKYKLTVAENPACPAVPSAAYVPSERKDNQST         | 1575 |
| QY     | 891   | LMPVDSGPEPESIRNTRSRSTPAIRASDPATCP,IPGCETIETISKGTGLGLSTVGS      | 950  |
| Db     | 1576  | PAYV---APLEPPIPSRSRSTPAVFAADPATCP,IIIPGCETIETISKGTGLGLSTVGS    | 1632 |
| QY     | 951   | DTLLGAFINHYVEEGAACADGRLMAGDQILEVNGIDLRKATHDAINVLROTPOVRVLT     | 1010 |
| Db     | 1633  | DTLLGAFINHYVEEGAACADGRLMAGDQILEVNGIDLRKATHDAINVLROTPOVRVLT     | 1692 |
| QY     | 1011  | LYRDEAPYKKEEVEDTLTIELOKKPGKGLSTVGRKNDGVFVSDIYKGGIADPDGRLI      | 1070 |
| Db     | 1693  | LYRDEAPYKKEEVEDTLTIELOKKPGKGLSTVGRKNDGVFVSDIYKGGIADPDGRLI      | 1752 |
| QY     | 1071  | OGDOILLVNGEDVARNASQEAVALLCSG,GTVTLLEVGRITKAGPFHSERRPSQTSQVSEGS | 1130 |
| Db     | 1753  | OGDOILLVNGEDVARNASQEAVALLCSG,GTVTLLEVGRITKAGPFHSERRPSQTSQVSEGS | 1812 |
| QY     | 1131  | LSSTFPRLSGSTSESSLESSKKNALASLFGKRTYEMKKKGTDLGSLIAGCVSSPLGD      | 1190 |
| Db     | 1813  | LSSTFPRLSGSTSESSLESSKKNALASLFGKRTYEMKKKGTDLGSLIAGCVSSPLGD      | 1872 |
| QY     | 1191  | VPIFIAMNHPRGVAAQCTOKRVDRIYVTCIGSTSEBEMTHTQAVNLILKNASGSIEMQVAG  | 1250 |
| Db     | 1873  | VPIFIAMNHPRGVAAQCTOKRVDRIYVTCIGSTSEBEMTHTQAVNLILKNASGSIEMQVAG  | 1932 |
| QY     | 1251  | GDVSVTGHNOEPASSLSLSTGLTSTISIFODD,IGRPQCKSTLERGPGLGFSIVGIGYS    | 1310 |
| Db     | 1933  | GDVSVTGHNOEPASSLSLSTGLTSTISIFODD,IGRPQCKSTLERGPGLGFSIVGIGYS    | 1992 |
| QY     | 1311  | PHGGLPIYKTVFRAKGAASDGRKLKGGDQILVANGOSLEGYTHREAVAILKRTGTYTLM    | 1370 |
| Db     | 1993  | PHGGLPIYKTVFRAKGAASDGRKLKGGDQILVANGOSLEGYTHREAVAILKRTGTYTLM    | 2052 |
| QY     | 1371  | VLS 1373   |      |
| Db     | 2053  | VLS 2055   |      |
| RESULT | 4   |  |      |
| ID     | 008783  | PRELIMINARY; PRT; 526 AA.                                      |      |
| AC     | 008783;   |  |      |
| DT     | 01-JUL-1997 (TrEMBLrel. 04, Created)                |  |      |
| DT     | 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)   |  |      |
| DT     | 01-MAR-2001 (TrEMBLrel. 16, Last annotation update) |  |      |
| DE     | 90KF BINDING PROTEIN 1 (FRAGMENT).                  |  |      |

| Query Match  | Best Local Similarity      | Matches | Conservative | Mismatches | Indels | Gaps |
|--|----------------------------|---------|--------------|------------|--------|------|
| 32.9%; Score 2397.5; DB 11; Length 526;                                  | 86.6%; Pred. No. 2,8e-129; | 457;    | 34;          | 32;        | 5;     | 2    |
| 98p-1.   |                            |         |              |            |        |      |
| GN Mus musculus (Mouse).   |                            |         |              |            |        |      |
| OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;     |                            |         |              |            |        |      |
| OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.    |                            |         |              |            |        |      |
| OC NCBI_TaxID=10090;   |                            |         |              |            |        |      |
| OC (1)   |                            |         |              |            |        |      |
| RP SEQUENCE FROM N.A.  |                            |         |              |            |        |      |
| RC TISSUE=PANCREAS;  |                            |         |              |            |        |      |
| RA Lee S.S., Weiss R.S., Javier R.T.;                                    |                            |         |              |            |        |      |
| RL Submitted (Apr-1997) to the EMBL/GenBank/DBJ databases.               |                            |         |              |            |        |      |
| DR EMBL; AF000166; AAB57835.1; -.  |                            |         |              |            |        |      |
| DR HSSP; Q12959; IPDR.   |                            |         |              |            |        |      |
| DR InterPro; IPR001478; -.   |                            |         |              |            |        |      |
| DR Pfam; PF00595; PDZ; 4.  |                            |         |              |            |        |      |
| DR SMART; SM00228; PDZ; 1.   |                            |         |              |            |        |      |
| FT NON_TER   |                            |         |              |            |        |      |
| FT 1   |                            |         |              |            |        |      |
| SEQUENCE 526 AA; 54948 MW; 3843C5E3AA2410E8 CRC64;                       |                            |         |              |            |        |      |
| QY 848 FLSLTKTAMTYKLTITHANPDSQAVPSPAAGASGKKNSQSLMYPQSGSPPEPSIRNT 907     |                            |         |              |            |        |      |
| DB 2 FLSLTKTAAATVATVILVRAENPACPAPPAVSATVYSGSKNSQTPAPV--ADPLEIPST 58      |                            |         |              |            |        |      |
| QY 908 SRSSTPAIRASDPATPCPIIPGCEETTEIEIKGRYGLSLIVSGSDTLGAFIITHEVEEGA 967  |                            |         |              |            |        |      |
| DB 59 SRSSTPAVPAASDPAPCPPIIPGCEETTIGVSGGTGLSLIVSGSDTLGAFIITHEVEEGA 118   |                            |         |              |            |        |      |
| QY 968 ACKDGRIMAGDDILEVNGIDLRKATHDEALINLRQTPQVRYRLTYRDEAPYKEEYCDTL 1027  |                            |         |              |            |        |      |
| DB 119 ACKDGRIMAGDDILEVNGIDLRKATHDEALINLRQTPQVRYRLTYRDEAPYKEEYCDTF 178   |                            |         |              |            |        |      |
| QY 1028 TIE--LQKKPKGKGLSTIYGRKNDGVFVSDIYKGGIADPDGLIGDDILVNGEDVFN 1085    |                            |         |              |            |        |      |
| DB 179 TIEIQLQKRPKGGLSTIYGRKNDGVFVSDIYKGGIADPDGLIGDDILVNGEDVFN 238       |                            |         |              |            |        |      |
| QY 1086 ASQEAVALILKCSLGTVLEVGRIKAPPHSERRPSQVSQVSEGLSFTPLSGSSTSE 1145     |                            |         |              |            |        |      |
| DB 239 ATQEAVALILKCSLGTAVTLEVGRIKAPPHSERRPSQVSQVSEGLSFTPLSGSSTSE 298     |                            |         |              |            |        |      |
| QY 1146 SLESSSKNALSEIOGLTVEYMKKGRPDLSIGTAGVGSPLGADPRTIAMNHPGVNA 1205     |                            |         |              |            |        |      |
| DB 299 SLESSSKNALSEIORLTVEYMKGRPADSLISIGTAGVGSPLGADPRTIAMNHPGVNA 358     |                            |         |              |            |        |      |
| QY 1206 QTOKLRVGDRIYITIGTSTEGMTHTQAVMLLNASSGSIIMQVAVAGDVSVTGHNDEPAS 1265 |                            |         |              |            |        |      |
| DB 359 QTOKLRVGDRIYITIGTSTEGMTHTQAVMLLNASSGSIIMQVAVAGDVSVTGHNDEPAS 418   |                            |         |              |            |        |      |
| QY 1266 SLSLSTGLTSTISIFODDLGAPPOCKSTLERPDGLFSTIVGSGPHGDLPIYKVTPEAK 1325  |                            |         |              |            |        |      |
| DB 419 PCLATFTGLTSSSIFPDDELGPPOCKSTLITDRGPDGLFSTIVGSGPHGDLPIYKVTPEAK 478 |                            |         |              |            |        |      |
| QY 1326 GAASDGRILKRGDQIIAVNGSLSEGTHEEAVAILIKRTGTYTLWYLS 1373             |                            |         |              |            |        |      |
| DB 479 GAASDGRILKRGDQIIAVNGSLSEGTHEEAVAILIKRTGTYTLWYLS 526               |                            |         |              |            |        |      |
| RESULT 5   |                            |         |              |            |        |      |
| ID 043798 PRELIMINARY: PRT: 453 AA.                                      |                            |         |              |            |        |      |
| AC 043798:   |                            |         |              |            |        |      |
| DT 01-JUN-1998 (TREMBLrel. 06, Created)                                  |                            |         |              |            |        |      |
| DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)                     |                            |         |              |            |        |      |
| DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)                   |                            |         |              |            |        |      |
| DE MULTI PDZ DOMAIN PROTEIN 1 (FRAGMENT).                                |                            |         |              |            |        |      |
| GN MUPPL.  |                            |         |              |            |        |      |
| OS Homo sapiens (Human).   |                            |         |              |            |        |      |
| OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;     |                            |         |              |            |        |      |
| OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.           |                            |         |              |            |        |      |
| NCBI_TaxID=9606;   |                            |         |              |            |        |      |

[1]  
 RN SEQUENCE FROM N.A.  
 RP TISSUE-BRAIN  
 RX MEDLINE-98196865; PubMed-9537516;  
 RA Ulmer C., Schumck K., Fige A., Lubbert H.;  
 RT "Cloning and characterization of MUPP1, a novel PDZ domain protein.";   
 RL FEBS Lett. 424:63-68(1998).  
 DR EMBL: AJ001319; CAA04680.1;  
 DR HSP: Q12959; 1PDR.  
 DR InterPro: IPR001478;  
 DR Pfam: PF00595; PDZ; 4.  
 DR SMART: SM00228; PDZ; 1.  
 FT NON\_TER  
 SO SEQUENCE 453 AA; 47423 MW; DA7915AD493C5DB8 CRC64;

Query Match 32.2%; Score 2252; DB 4; Length 453;  
 Best Local Similarity 98.2%; Pred. No. 1.1e-126;  
 Matches 445; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

OY 921 ATCPICPCETIETISKRTGLISVSGSDTLGAFIIHEYEGGACKGRMAGDOI 980  
 DB 1 ATCPICPCETIETISKRTGLISVSGSDTLGAFIIHEYEGGACKGRMAGDOI 60  
 OY 981 LEVNGIDLRKATHDAINVLKOTPORVRLTYRDEAPYKEEPCDTLITELOKKPGKIG 1040  
 DB 61 LEVNGIDLRKATHDAINVLKOTPORVRLTYRDEAPYKEEPCDTLITELOKKPGKIG 120  
 OY 1041 LSIYGRNDTVGVFSDIVKGIADPDGRLIQDQILVNGEDVRNASEVAALLKCSLG 1100  
 DB 121 LSIYGRNDTVGVFSDIVKGIADPDGRLIQDQILVNGEDVRNASEVAALLKCSLG 180  
 OY 1101 TTTLEVGRTKAGPHERPSPQTSVSEGSLSFTPLSGSTSSLSKKNALASFI 1160  
 DB 181 TTTLEVGRTKAGPHERPSPQTSVSEGSLSFTPLSGSTSSLSKKNALASFI 240  
 OY 1161 QGLRTVEKKKGTDSLIGISAGVSPGLDVPITFAMMHPTEVAQOTKLRNGDRIYTC 1220  
 DB 241 QGLRTVEKKKGTDSLIGISAGVSPGLDVPITFAMMHPTEVAQOTKLRNGDRIYTC 300  
 OY 1221 GTSIEGKTHQAVNLKKNASGSIEMQVAGDVSVTGHQEPASSLSFTGTSSTIFQ 1280  
 DB 301 GTSIEGKTHQAVNLKKNASGSIEMQVAGDVSVTGHQEPASSLSFTGTSSTIFQ 360  
 OY 1281 DDLGPPCKSTILEGPGIGSIYGVSPGDIPIYKTFEAGASSEGRRLRGQI 1340  
 DB 361 DDLGPPCKSTILEGPGIGSIYGVSPGDIPIYKTFEAGASSEGRRLRGQI 420  
 OY 1341 IAVNGQSLGVTHHEVAVALKRTKGTVMVLVS 1373  
 DB 421 IAVNGQSLGVTHHEVAVALKRTKGTVMVLVS 453

RESULT 6  
 O9H3N9 PRELIMINARY; PRT; 1134 AA.  
 AC O9H3N9:  
 DT 01-MAR-2001 (Tremblrel. 16; Created)  
 DT 01-MAR-2001 (Tremblrel. 16; Last sequence update)  
 DT 01-MAR-2001 (Tremblrel. 16; Last annotation update)  
 DE PDZ DOMAIN PROTEIN 3' VARIANT 4.  
 GN HINADL.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Soejima H., Kawamoto S., Akai J., Miyoshi O., Arai Y., Nishikawa N.,  
 RA Kimura A., Okubo K., Mukai T.;  
 RT "Cloning, characterization and chromosomal mapping of the two novel  
 RT heart specific genes."  
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL: AB044807; BAB19683.1;  
 SO SEQUENCE 1134 AA; 125244 MW; FB6C3DA92D7CER86 CRC64;

Query Match 26.0%; Score 1820.5; DB 4; Length 1134;  
 Best Local Similarity 39.3%; Pred. No. 3.4e-100;  
 Matches 474; Conservative 140; Mismatches 310; Indels 281; Gaps 36;

OY 1 MYCCRRVPPPTQSELDSLIDCDITELTEKPHVLDGEFTGSETEDPVYLAAMDAGOSTEBV 60  
 DB 92 LVCCRRLE--DDEASVDEPRRTETSLPE--TEVDHMDVDNTEEDD----- 133  
 OY 61 QAPLAMEAGIOHTELEKSGKGLFSIIDYODPIDPASTVIIRSLVPGIAEKDGRLLP 120  
 DB 134 -GELALMSPEYKIVELVDCGKGLFSIIDYODPIDPASTVIIRSLVADGVAERSGGLP 192  
 OY 121 GDRLMFVNDVNLNLSLEAEALKAPSGVIRIGAVKAPLPSPEG-----YVSAKED 174  
 DB 193 GDRLVSNVEYCLDNTSLAEAVELKAVPPGLVHIDICKPLVEDNEEESCYILHSSNEDK 252  
 OY 175 -SFLYPPHSCBA-----GLADKPLRADIALVGTMDADLVDESTESPSPENDSIY 226  
 DB 253 TFSCTIHDINSLILEAPKGRFDEPYEK-----EELVDE-----PFLDGKSPH 297  
 OY 227 STQASITLSLHSSCGDGLNYGSSLPSPKDYIENSCDPVLDLHMSLELYQNLEBOD 286  
 DB 238 SQOKET-----EOKKAME-----MHEFLTPLOENDE 325  
 OY 287 E-----NTPSYDISMGPASGFTINDYTPRANAIEQDYECENTIVWTESHPSEV 334  
 DB 326 EREMVLDEBELYQDPSMELYPLS--HIGATVPVSN--ELHFGQMLHNDPSPS 380  
 OY 335 ISSAELPSVLPDSACKGSEHLEQSSSLACMACVMAKQVSKSFERTINAKGNSLSMT 394  
 DB 381 QEARTGRVYSGEA-----QPYGYC--PENYKKNFVMSLP--SVPTSGNSQOGR 428  
 OY 395 VSANRGLGMIYRSIIHGAIISDRIRIAGDIIISNESTISVTNAQARALRHSLSIG 454  
 DB 429 -FDDLENLSLAKTSLDLMITND-----VQG 454  
 OY 455 PDIKITY--VPAHELEPKISLQO--OSGRVMAIDIFSSTG-----RDIPLEPERGE 505  
 DB 455 PSLIDLPLVQARRQEDLPYQOATRVISK--ASATGMLSSHYADTCELPEREGE 512  
 OY 506 GESELOMTATYNNMQPRVELMREPSKSLGSIYGGKMGKRLNGEVMRIFIKHYLE 565  
 DB 513 GEET---PNSHMCPPRIVELIFREPNSVLGISIVGGQTVIKRLKNGELKGIETKOYLE 568  
 OY 566 DSPACKNGTLKPGDRIVEA-----PSQSESEPEKA-----PLCSVPPPPSA 607  
 DB 569 DSPACKTNALKTGDKILEVSGVDLONASHSEAVEAIKKNAGNPVYFIVOSLSTPRVIRNV 628  
 OY 608 FAEWG-----SDHTQSSASK-----ISQDVX--EDEFGYSMKNIER 643  
 DB 629 HNKANKITGNQNOQOEKKEKRGQTAPPKMLPPYKALKALTDSDSDEDEADFTDKIKOR 688  
 OY 644 YCTTLGELHMTLEKSGHGLSLAGNDRSMSPFIYICIDPONGANGADGRQIDELLE 703  
 DB 669 YADLGEHIIIELEKDKNGSLAGNDRSMSPFIYVGINEGPAADGRIDELLE 748  
 OY 704 INGOILYGRSHONASSIITKCAPSKYKIFIRKDAVNOMAVCPGNAVEPLPSSENLOK 763  
 DB 749 INNOLLYGRSHONASSIITKAPSKYKLVFIRNEDAVNOMAYP-----PPVPSSP----- 799  
 OY 764 ETEPTVTSDAAVDSLFRKNVOHLELPKQDGLGIAISEEDTLGVIILKSLTEHGAATD 823  
 DB 800 -----SSI-----EDQSGTEPISSE-----D 816  
 OY 824 GRLKYGDQILAVDDIIVGYPIEKFISLTKTKMYKLTIIHENDSOAVPSAAGASGE 883  
 DB 817 GLEVGIRQLPSESFKLA-----VSQMKQOKYPTKVSF-----SSOEIPLAPSS--- 862  
 OY 884 KNSQSILMVPOGSPPEPESIRNTSRSSPTAIFADSPATCPIIPCETIETISKRTGLG 943

```

Db 863 -YHSTDADFTGCGFOAPLSV-----DPATCPYVGOEMITEISGRSGIG 907
OY 944 LSTVGGSDTLTGAFITHEVEYEEGAACKDGRMLAGDQILEVNGIDLRKATHEALNVLRQT 1003
Db 908 LSTVGGSDTLPLNATVITHEVEYEEGAARDGRMLAGDQILEVNGVLRNRSSEALITALRQT 967
OY 1004 PORVRLTLVDEAPYKEEVEDTLTTELQKKPGKGLSLVGRKNDGVFVSDIVKGGIA 1063
Db 968 POKRLVLYVREAHYRDEENLEIFPVLDQKKAGKGLSLVGRKNGSGVFTISDLVKGGAA 1027
OY 1064 DPGRLLIQQDQIILVNGEDVNRASQEAVALKCSLGTVTLVGRIRKAGPFRSPSQT 1123
Db 1028 DLDRLIQQDQIILVNGEDVNRASQEAVALKCSLGTVTLVGRIRKAGPFRSPSQT 1087
OY 1124 SOVSE 1128
Db 1088 SOSAE 1092

RESULT 7
060833 PRELIMINARY; PRT: 1552 AA.
AC 060833;
DT 01-AUG-1998 (TEMBLrel. 07, Created)
DT 01-AUG-1998 (TEMBLrel. 07, Last sequence update)
DT 01-MAR-2001 (TEMBLrel. 16, Last annotation update)
DE INADL, C-TERM VARIANT2.
GN INADL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homindae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Philipp S.;
RL Submitted (MAR-1998) to the EMBL/GenBank/DDBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE-97424368; PubMed-9280290;
RA Philipp S., Flockerzi V.;
RT "Molecular characterization of a novel human PDZ domain protein with
RT homology to INAD from Drosophila melanogaster.";
RL FEBS Lett. 413:243-248(1997).
DR EMBL; AJ224747; CA12112.1; -.
DR HSSP; P31016; 1BEG.
DR InterPro; IPR001478; -.
DR Pfam; PF00595; PDZ; 8.
DR SMART; SM00228; PDZ; 1.
DR SRA; SRA00228; PDZ; 1.
SQ SEQUENCE 1552 AA; 170322 MW; FF80D9A65350F996 CRC64;

```

Query Match 22.2%; Score 1550.5; DB 4; Length 1552;  
 Best Local Similarity 37.3%; Pred. No. 8e-84;  
 Matches 419; Conservative 131; Mismatches 293; Indels 281; Gaps 36;

```

Db 839 SOKKEI-----EOKSAWE-----KHEFLTRLOQWDE 866
OY 287 E-----NTPSVDISMGPAAGFTINDYIPANAIQOYCEMTIWTESHLESEV 334
Db 867 EREMLVDEEYELVODPSPSMELPLS--HIOEATPVSVN--ELHFGTOMLHNEPSES 921
OY 335 ISSAELPSVLPDSAGKSEHLEBQSSSLACNAECVMLONVSKESPRTINAKNGSGMT 394
Db 922 QEATGRTVYSGEA-----QPYGVC--PENMKENFWESLP--SVPTSGSQGR-- 969
OY 395 VSANKDGLMIVRSIHGGAISRDRGRIAGDCILISINEESTISVTNAQARMLRHSIG 454
Db 970 -FDLEMLNLSIAKTSLDLGMIPND-----VQG 995
OY 455 PDIKITY-VPAEHLEEFKISLQ--QSGRVNADLFSSYTG-----RDIPELPEREGE 505
Db 996 PSLIIDLPPVAAQRREQEDLPYHOATRVISK--ASAVTGLMSRYATDCELPEREGE 1053
OY 506 GESELONTAVSNMOPRVELMREPESKLSISTVGRGMSRLSNGEVNRGIFRHYLE 565
Db 1054 GEET-----PNFSHWGPRRIYEIFREPVSIGISIVGGQYVYKRLKNEELKGIKYLE 1109
OY 566 DSPAGKNGTLKPGDRIYEA-----PSQSESEPEKA-----PLCSVPPPPSA 607
Db 1110 DSPAGKTNALKTGDKILEVSGVDLQNSHSEAVAIKKNAGNPVYFIVQSLSSTPRVIVN 1169
OY 608 FAENG-----SDHTQSSASK-----ISQDVUK--DDEGYSKKNRER 643
Db 1170 HNKNKRTTGNONDOTQKKERQGTAPPMKLPYRYALYDDSDENEEDAFDQKTROR 1229
OY 644 YGTLTGELHMELEKSGHGLSLAGNKRDSRMSVFYIGDIPNGAAGDRLQIADLELLE 703
Db 1230 YADLPGLHIIIELEKDKNGGLSLAGNKRDSRMSVFYIGDIPNGAAGDRLQIADLELLE 1289
OY 704 INGOILYGRSHONASSIITKAPSKVKIIFIRNKDAYQMAVCPGNAVEPLPSNSENLOK 763
Db 1290 INNOILYGRSHONASSIITKAPSKVKIIFIRNKDAYQMAVCPGNAVEPLPSNSENLOK 1340
OY 764 ETEPTVTSAAVDSLFSKVNQHLELPKDDGGLGIALSEEDTLSGVILKLTENVAATD 823
Db 1341 -----SST-----EDSGTEPISSSE-----D 1357
OY 824 GRUKVGDQILAVDDEIVGVPIEKFISLTKAKTVLTTIHAENPDSQAVPSAAGASGE 883
Db 1358 GSELVGIIKQLPSESEFKLA-----VSQMKQOKPTIVSF-----SSQELPLAPASS--- 1403
OY 884 KKNSSQSLMVPQSGSPESIRNTSRSTPAIFASDPATCPITIPGCTTETIGKRTGIG 943
Db 1404 -YHSTDADFTGCGFOAPLSV-----DPATCPYVGOEMITEISGRSGIG 1448
OY 944 LSTVGGSDTLTGAFITHEVEYEEGAACKDGRMLAGDQILEVNGIDLRKATHEALNVLRQT 1003
Db 1449 LSTVGGSDTLPLNATVITHEVEYEEGAARDGRMLAGDQILEVNGVLRNRSSEALITALRQT 1508
OY 1004 PORVRLTLVDEAPYKEEVEDTLTTELQKKPGKGLSLVGRK 1047
Db 1509 POKRLVLYVREAHYRDEENLEIFPVLDQKKAGKGLSLVGRK 1552

RESULT 8
043742 PRELIMINARY; PRT: 1582 AA.
AC 043742;
DT 01-JUN-1998 (TEMBLrel. 06, Created)
DT 01-JUN-1998 (TEMBLrel. 06, Last sequence update)
DT 01-MAR-2001 (TEMBLrel. 16, Last annotation update)
DE INADL, C-TERM VARIANT3.
GN INADL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homindae; Homo.
OX NCBI_TaxId=9606;

```

```
D0      11 0  RNNANNLIGUNODIQENKRNKGIAPEPMIPPEPTIKALIDUSDENEEDATFOQINQR   12
Oy      64d YGTLTGELHMTLELKGHSGLSLAGNKRDRSRMSVFTVGIDPNGAAGKDGRLQIADLELE 703
```

QY 758 ENLONKETEPTVTTSDAANDLSFKNVQHLPLP-KDOGGGLAISEEDPTLSGVIIKSLTE 816  
117 -----FPLSLSPSPVEDLGGTELYSSEEE--SSVDAKHUPE 151



|        |  |   |               |
|--------|--|---|---------------|
| Oy     | 817  | HGVATGCRRLKVGQIIILAVDEIVYGVPIKFIISLTKTMMYKILIHENEDSOAVPSA   | 876           |
| Db     | 152  | PESS-----KPELDSQVVDNNKAAEQKE-----                           | 187           |
| Oy     | 877  | AGAAAGEKKNSSQSLMWPQSGSEPEESIRNTSS-----                      | 927           |
| Db     | 188  | QOFTVSTGVSSSD--SPSSPAPLCSAHADVSGNFOAPLVPDPAPLSPATPCPIVP     | 245           |
| Oy     | 928  | GCETTIEISKRTGLGISTYVGSPTLLGAFITHEVYEEGACKDGLNMGDDITLEVNGID  | 987           |
| Db     | 246  | GOEMIEISKRSGLGISTYVGSPTLDAIVTHEVYEEGAARDGLNMGDDITLEVNGVD    | 305           |
| Oy     | 988  | LKRAITHEAIVLVTQTPQVRVRLTYLDEAPYKEEVECDPLTTEILOKPKGGLSTVGR   | 1047          |
| Db     | 306  | LRRSSHEAIVLALQTPQKRVLYVREAOYRENELEFVLVDLCKITGRGLSTVGR       | 365           |
| Oy     | 1048   | NDTGVPVSDIYVKGSIADPDGRILQDGIILVNGEDVRNASSOEAAVALLKSLGVTLEV  | 1107          |
| Db     | 366  | SGSGVFETSDIYKCGAADLDRLIRGDILSVNEEDNRHNSOETVATILKCVQGLVLEIG  | 425           |
| Oy     | 1108   | RKKAGPFSERPEPOTSOVSEGLSSTPTPLGSSSTSELES-----                | 1166          |
| Db     | 426  | RLRGSWMAASKTSQNSQGDHSHHSCRP-SFAPVITSLONLVGTKRSSDPQKCTE-E    | 483           |
| Oy     | 1162   | GLRTVEKKKGFPTDSLIGISIVAGVSGPLGDPYFIAMNHPGVAADQTKLWDGRIVTTCG | 1221          |
| Db     | 484  | EPRTVEIRRELSDALGISIVGSKGSPGLDPIPIAMIQNGVAARQTKLWDGRIVTSING  | 543           |
| Oy     | 1222   | TSTGCMHTQAVNLTKMASSISEMNOVAGDGVVYVGHIOEPASSLSSTGLTSTISIOD   | 1281          |
| Db     | 544  | OPLOGLSITDVLNLLTKMAFRITLIDVAVDNTISAIATOLEIMSAGS-----        | 589           |
| Oy     | 1282   | DLGPPCKSTYLERGPD  | 1298          |
| Db     | 590  | QLGSP-----TADNRPE   | 601           |
| RESULT | 10   |   |               |
| ID     | 015249   | PRELIMINARY;  | PRT; 1524 AA. |
| AC     | 015249;  |   |               |
| DT     | 01-JAN-1998 (TREMBlrel. 05, Created)                                 |   |               |
| DT     | 01-JAN-1998 (TREMBlrel. 05, Last sequence update)                    |   |               |
| DT     | 01-MAR-2001 (TREMBlrel. 16, Last annotation update)                  |   |               |
| DE     | PDZ DOMAIN PROTEIN.  |   |               |
| OS     | Homo sapiens (Human).  |   |               |
| OC     | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;    |   |               |
| OC     | Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.            |   |               |
| OX     | NCBI_TaxID=9606;   |   |               |
| RN     | [1]  |   |               |
| RP     | SEQUENCE FROM N.A.   |   |               |
| RX     | MEDLINE=97424368; Pubmed=9280290;                                    |   |               |
| RA     | Phillip S., Flockezy V.;   |   |               |
| RT     | "Molecular characterization of a novel human PDZ domain protein with |   |               |
| RT     | homology to IMD from Drosophila melanogaster.";                      |   |               |
| RL     | FEBS Lett. 413:243-248(1997).  |   |               |
| RN     | [2]  |   |               |
| RP     | SEQUENCE FROM N.A.   |   |               |
| RX     | MEDLINE=96224170; Pubmed=8617505;                                    |   |               |
| RA     | Lenon G., Aufreay C., Polymeropoulos M., Soares M.B.;                |   |               |
| RT     | "The I.M.A.G.E. Consortium: an integrated molecular analysis of      |   |               |
| RT     | genomes and their expression.";                                      |   |               |
| RL     | Genomics 33:151-152(1996).   |   |               |
| DR     | EMBL; AJ001306; CAA04666.1; -.                                       |   |               |
| DR     | HSSP; P31016; 1BFE.  |   |               |
| DR     | InterPro; IPR001478; -.  |   |               |
| DR     | Pfam; PF00595; PDZ; 8.   |   |               |
| DR     | SMART; SM00228; PDZ; 1.  |   |               |
| DO     | SEQUENCE 1524 AA; 167323 MW; E3BBA7EC856A954D CRC64;                 |   |               |

|             |   |
|-------------|---|
| Query Match | 19.98; Score 1390.5; DB 4; Length 1524; |
|-------------|---|

[illegible]



Db 1449 LSIYVCKDPTL-----VNCVDLNRSSHEEATPALROT 1480  
 Qy 1004 PQRVRLTLYRDEAPYKEEVECDTLIELQKRGKGLSTYVKR 1047  
 Db 1481 PQRVRLTLYRDEAPYKEEVECDTLIELQKRGKGLSTYVKR 1524

## RESULT 11

009515 PRELIMINARY: PRI: 2208 AA.  
 AC 009515:  
 DT 01-JUN-1998 (TREMBlrel. 06, Created)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
 DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)  
 DE HYPOTHEtical 240.1 KDA PROTEIN C52A11.4 IN CHROMOSOME II.  
 GN C52A11.4.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditidae; Rhabditoidea;  
 OC Rhabditidae; Pelodierinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL NZ;  
 RA Coles L., Sulston J.;  
 RL Submitted (DEC-1998) to the EMBL/Genbank/DBJ databases.  
 DR EMBL; 246792; CA86769.2; JOINED.  
 DR EMBL; 246795; CA86789.2; JOINED.  
 DR EMBL; 246792; CA86789.2; JOINED.  
 DR HSSP; Q12959; 1PDR.  
 DR InterPro: IPR001478; -  
 DR InterPro: IPR002173; -  
 DR Pfam: PF00595; PDZ; 15.  
 DR PROSITE; PS00584; PKB\_KINASES\_2; UNKNOWN\_1.  
 DR PROSITE; PS00584; PKB\_KINASES\_2; UNKNOWN\_1.  
 SQ SEQUENCE 2208 AA; 240120 MW; B6894F8D603E79FB CRC64;

Query Match 14.9%; Score 1044.5; DB 5; Length 2208;  
 Best Local Similarity 21.4%; Pred. No. 2.6e-53;  
 Matches 385; Conservative 176; Mismatches 324; Indels 911; Gaps 40;

Qy 64 LAMVAGIOHIELEKSGKGFSLDYODPDIPASTVIRSLVGVGIAENKGRLLPGDR 123  
 Db 703 LAMVAGIOHIELEKSGKGFSLDYODPDIPASTVIRSLVGVGIAENKGRLLPGDR 762  
 Qy 124 LMFVNDVNLNSLEAEVAKLAPSGYRIVAKPLPSPEEGYSAKEDSFLYPHSC 183  
 Db 763 LMFVNDVNLNSLEAEVAKLAPSGYRIVAKPLPSPEEGYSAKEDSFLYPHSC 819  
 Qy 164 EAGGLAD-----KPLRADLAVGTNDADVDDESTFE-SPLSPENDSTYTOASTL 233  
 Db 820 KICGLVQLNNIESFLIKALIKOEVR--QOHEVNGVSRSEFSMTSPSCSTRSVSPGSP 877  
 Qy 234 SLHGSSCGDLNYSGLSPSPPKVYENSCDVLDLHMSLEELYQNLERODEMTPSVD 293  
 Db 878 SLHGSSCGDLNYSGLSPSPPKVYENSCDVLDLHMSLEELYQNLERODEMTPSVD 886  
 Qy 294 ISMGASGFTINDYPPANAIEQOYECENTIVWTESHLPSEVISSAELPSVLDGAGKGE 353  
 Db 887 ISMGASGFTINDYPPANAIEQOYECENTIVWTESHLPSEVISSAELPSVLDGAGKGE 894  
 Qy 354 HLEBOSSLAECVNLQNVSKESFEFTINIKAGNSSLGMYTSANKD-GL-GMYRSTIH 411  
 Db 895 HLEBOSSLAECVNLQNVSKESFEFTINIKAGNSSLGMYTSANKD-GL-GMYRSTIH 929  
 Qy 412 GAISRDGIAIGDCILSINESTISVTNAQARMLRSLIGPDIKITYP----- 463  
 Db 930 KKAVALDRIQVDFITKINTESLANTNSQARALILKRNINLVGTICNTYITISDAKTKR 989  
 Qy 464 -----AEH 466  
 Db 990 ERFQRPSESSPIINRLSPKVPKFRSPFMQROESQSKTEMTDDETPASIMTDSMSEH 1049

Qy 467 LEEFKISLG-----QSGRYVAL----- 484  
 Db 1050 IKTFLDAGSSRSHHDEQEVNRMSRLIDGYEVDDVNLIIKEATDTIELRVLHKTKD 1109  
 Qy 485 -----DIRSS-----YTGRDI----- 495  
 Db 1110 WSNCHKREERLESPLPPPPREVLSPSPKSPVASVOPTPRDLBEVLRSTASSLEYHSG 1169  
 Qy 496 -----PELPEREGE----- 505  
 Db 1170 ORTSQHLILSTEEVLOATPPSSPENKSEVPPSPISPGIKLAGEVTAPEETIEVYKQAE 1229  
 Qy 506 ----- 505  
 Db 1230 NVDRAEATAGAEVATSTPTPAEALGNSKDESTTSSQSVGLQTOALNSTEVNSMS 1289  
 Qy 506 -----GESELONTA-----YSNNQPRVELMREPSKSLGISYVGR---GMGS 547  
 Db 1290 RVTSPRTSGSES-LQNGARQLVRKSYWGCEARTVTLVREPKNSFGISYGVESYQKGG 1348  
 Qy 548 RLSNCEVARGITIKIVYLEDSPAGKNGITLPGDRIVEAP----- 585  
 Db 1349 LPGTNTVCGIFIKSVLPNSPAGRSQNMNGDRVTSVNDVLDRLATHEQAVNAIKNASNP 1408  
 Qy 586 -----SQSESEPKAPLCSVPP----- 602  
 Db 1409 VAFVLQSLHTNQNNINSASNTGVSVRFENAKPPEELPTALVTPPLAMISSGSSSTSK 1468  
 Qy 603 -----PPPS---APAEKSGDHTQSSASKISQDVKED----- 631  
 Db 1469 PAMNPPPSISYTTTTSMESEKKEEATSSPDIGRENTVYKRSMDVEREQEAPNVEET 1528  
 Qy 632 ----- 631  
 Db 1529 SIYKERTPPRKISAKSNSQKESNDRKEIKOKSVREMSVESKKSVSRIKKHROESTII 1588  
 Qy 632 ----- 631  
 Db 1589 KKSPPNETAPLIVSDVSSSETHEDPEQAMSPSTPDTADAMRALGIDDDSAARQIKND 1648  
 Qy 632 -----EFGYSWKNIERYGTTLGELHMI-ELEKSHGSLGLSLAGNKRDRMSVYIGIDPN 686  
 Db 1649 GEPSKFEFTYACKIRKXKSDSGDELVLVACERPDGGLISLAGNKRDRMSVYIGIDPN 1708  
 Qy 687 GAAGKDGRLQIDELLEINGQILYGRSHONASSITK-CAP--SKKTIIFIRKADVNOQA 743  
 Db 1709 CPLA-----IRPDDELLEINGRLINKISHVAASVARECCDOHONTEIYLRNRGALKSTA 1764  
 Qy 744 VCPGNAVEPLPSSENLOKTEPEVTYTTSDAVIDLSFKNVQHLPLPDQGLGIAISEE 803  
 Db 1765 -----POMPTTGTGTEH-----AAPNQ 1781  
 Qy 804 DTLGVIKISLSEHGAATDGRKLVGDQILAVDDEIVVGYPIEKFTSLTKAKMTVKLTI 863  
 Db 1782 EELS----- 1785  
 Qy 864 HAENPDQAAPSAAGAAGEKKNSSQSLAVPOGSGPEPEESINTSRSSSTPAIFASDPATC 923  
 Db 1786 -----RKRSFSOE-----RQAIIN----- 1800  
 Qy 924 PIIPGCEITIEISKRTGGLSIVGSDTLGAFIIEHVEEGACAKGRMLAGDQILEV 983  
 Db 1801 -----GRETWIEIDKQKGLGLSIVGADVDLGTAVIHEVYSKGAABDRLPGQOVLEV 1856  
 Qy 984 NGIDLKRAKATHEAIVNLKQTPQVRRLTYRDE---APYKEEVECDTLIELQKRGKGL 1039  
 Db 1857 NGTSLRGVTHDSIAYLRTPRKVLILYRDVNDLSLDLPQIINIFEIDLYKTKTGGL 1916  
 Qy 1040 GLSIYGRKNDGVYVSDIYKGIADPGRLIGDQILLVNGEDVNASQEAVALKCSL 1099  
 Db 1917 GISIVGRKNEPVYVSEIYKGLAESDGLMTGDQILEVNGKDVGCQOEDVAAALAKTKYT 1976  
 Qy 1100 GTVLELVEGRIRKAGPF----- 1114

```

Db 1977 GKVALKTEENNNDPVRVAVAVASSPAATPTSTASVYVTSAPATSVPEEDNVVEPIPGT5 2036
Oy 1115 -----HEKRRPSQTSOYSEGSLSFF-----PLS 1139
Db 2037 AREERNVPPAPPMREIITHTSPECCEFOQEPAGLSPTEPESSGCGFM5VOEEBQPT 2096
Oy 1140 GSSSTSESSSKKNALASPI-----OGLR--VEKMGPTDSIGIAGVSPGLDV 1191
Db 2097 PTTT-----SSNNNNNSLAIDIHDLKEESDPLVLEKRVYDQGLM-----GIGK--RSR 2147
Oy 1192 PPIAMHPPGVAAOQOKLAVGDRIVTCGTSEGTHTQAVNLKNAGSIEMOV 1247
Db 2148 GILVTSLOPSSAAA--EKIKVGRIILAV--NALPVSDQLSAVTFVKAGSGRLYLDI 2199

RESULT 12.
O70263 PRELIMINARY: PRT: 728 AA.
AC 070263; 070264;
AD 01-AUG-1998 (Tremblrel, 07, Created)
DT 01-AUG-1998 (Tremblrel, 07, Last sequence update)
DT 01-MAR-2001 (Tremblrel, 16, Last annotation update)
DE NUMB-BINDING PROTEIN LNX (LIGAND OF NUMB-PROTEIN X).
GN LNX.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN 11
RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
RC TISSUE-BRAIN, AND EMBRYO.
RX MEDLINE=98204916; PubMed=9535908;
RA Dho S.E., Jacob S., Wolting C.D., French M.B., Rohrnschneider L.R.,
RA MGladie C.J.;
RT "The mammalian numb phosphotyrosine-binding domain. Characterization
RT of binding specificity and identification of a novel PDZ domain-
RT containing numb binding protein, LNX."
RT J Biol Chem. 273:9179-9187(1998)
CC -1- FUNCTION: INTERACTS WITH THE PHOSPHOTYROSINE INTERACTION DOMAIN OF
CC THE NUMB PROTEIN.
CC -1- ALTERNATIVE PRODUCTS: TWO ISOFORMS; ISOFORM LNX (OR LNXP80) (SHOWN
CC HERE) AND ISOFORM LNX-B (OR LNXP70); ARE PRODUCED BY ALTERNATIVE
CC SPLICING.
CC -1- TISSUE SPECIFICITY: BOTH ISOFORMS ARE EXPRESSED IN THE HEART.
CC ISOFORM LNX IS ALSO EXPRESSED IN KIDNEY, LUNG AND SKELETAL MUSCLE
CC WHILE ISOFORM LNX-B IS ALSO EXPRESSED IN BRAIN.
CC -1- SIMILARITY: CONTAINS 4 PDZ DOMAINS.
DR EMBL: AF034745; AAC40075.1; -
DR EMBL: AF034746; AAC40076.1; -
DR HSSP: P31016; 1BFE.
DR MGD: MGI:1278335; Lnx.
DR InterPro: IPR001478; -.
DR Pfam: PF000957; zf-C3HC4; 1.
DR Pfam: PF00595; PDZ; 4.
DR SMART: SM00228; PDZ; 1.
DR Alternative splicing.
KW DOMAIN
FT DOMAIN 278 362 PDZ.
FT DOMAIN 385 467 PDZ.
FT DOMAIN 508 593 PDZ.
FT DOMAIN 638 723 PDZ.
FT MOTAGEN 188 188 Y->A: ABOLISHES BINDING TO NUMB PROTEIN.
FT MOTAGEN 188 188 Y->F: NO EFFECT ON BINDING TO NUMB
FT MOTAGEN 188 188 PROTEIN.
FT VARSPLIC 1 131 NUMBPLADDPSPSPPLCTVCGNHSHPENHFTYVEDVD
FT DILGICQALLDPLDPGCGTYCTCLITLNFVLEKDCPD
FT KPELVQCKRSSTLVNKLKLVTCPFTFDECEVLQRD
FT LQHHQTS -> MALLLLVLPWLSPPANYIDWNGNHLFYS
FT EL (IN ISOFORM LNX-B).
SO SEQUENCE 728 AA; 80156 MW; E2914BD364CCEC4 CRC64;
Query Match 8.48; Score 569; DB 11; Length 728;

```

Best Local Similarity 29.1%; Pred. No. 6.9e-27;  
Matches 188; Conservative 115; Mismatches 238; Indels 106; Gaps 23;

```
QY 779 SFFKVVHLELPD-----QGGL-GIAISEEDTSGVIITSLTEHGVAATDGRLLVGDQ 831  
      : : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |  
Db 130 TSCCKASHGILTKDKRRRSODCCPDGCASLMAITLS-----PEVSAAAT----- 173  
  
QY 832 ILAIVDEIYVGPIKEKTSILTAKMTVKLTTHAENPDSOAVP---SAAGAAGEKKSS 888  
      : : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |  
Db 174 ISLMTDEGLDP--AYYSVDGEPPVANSDSGSNTRAPREFRSMRSRKFKINRA 231  
  
QY 889 QS-LMWPGSGSEPEST---RNTSRSTPAIFASDPATPCIIPOCE-TTIEISKRTGL 942  
      : : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |  
Db 232 ISALRRRTSGSVVAHVADGGRDNSENTTYPEVF---PRFLFIDPGETISIKIRNPSE 288  
  
QY 943 GLSI--VGGSDTLCAFLIHVEYEGCAACKGRMLAGOLILEVNGIDRKATHDEAIIVL 1000  
      ||| ||||| : | : | : | : | : | : | : | : | : | : | : | : | : | : |  
Db 288 SLSTLVGGSERPLHIITQHTRGVARQRRLPGIITILKVNMDSNVPHTAVRL 348  
  
QY 1001 RCPQVRVLTRYRDE-----APYEEEVCDTLTTELQK-KFGCLGISYGRNDT 1050  
      : : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |  
Db 349 RQPCQVLRITYLREQKFRSRNAHPDSTGYPRDDSFHYLTKKSSPEDQLGIKLVRVDEP 408  
  
QY 1051 GFYFVDIYKGGIADPDGRLIGDOLLVNGEDEVRYNASDEAAALLKCSLGTYTEYGRIK 1110  
      : : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |  
Db 409 GFETFNVLNGVADRHGOELEENDRYALINGDLRGSPESAAHLIQASERNVHLVNS-- 466  
  
QY 1111 AGPFHSERRPS----QTQSYSEGSLSSFPTPLSGSTSESLESSSKKNALAASEIOGLRT 1165  
      : : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |  
Db 467 -----QVHQSSPDIDIQEAGWISNGQS----PGPER-----NTASKPAANCHE---KV 508  
  
QY 1166 VEMKGPPDLSLGISTAGVGPLGDPVPIFIAMNHPTGYAAOTOKLRVDRITYICGISTE 1225  
      : : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |  
Db 509 VSVMWDPSSELMTMGYGASAREMDLPITYVISVEGCVISRDKRTKDILLANGLIELT 568  
  
QY 1226 GUTHQAVALINKNAGSILEMVGAVAGDVSVYTGHHQEPASSLSPTGLTSTIFODDLCGP 1285  
      : : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |  
Db 569 EYSRENEVAAILKSASSVYLKAL---EK-----EQDAQEDCSPALADS-----NNHWTP 615  
  
QY 1286 P-----OQKSTITLEGPPDG-LGFSTYGVGYSFHEDLPITYVKTVPFAKG 1366  
      : : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |  
Db 616 PDMSPSWMWLDELPOLYLNCNDVLTIRNTGSLGFCFYGEVEXSGMKPKFISIVEGCT 675  
  
QY 1327 AASEGRLKRGDQILAVNGOSLEGVTHEEAVALIKRTKGTATLAWLS 1373  
      | ||| : | : | : | : | : | : | : | : | : | : | : | : | : | : |  
Db 676 PAYNDGRIRCGDILLAVNGRSTSGNIHCALARMLKELGRITLTIAS 722
```

RESULT 13  
Q23823 PRELIMINARY; PTR; 665 AA.  
AC Q23823:  
DT 01-NOV-1986 (TREMBLrel. 01, Created)  
DT 01-NOV-1986 (TREMBLrel. 01, Last sequence update)  
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
DE INMD PROTEIN.  
GN INMD.  
OS Calliphora vicina (Blue blowfly) (Calliphora erythrocephala).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Oestroidea; Calliphoridae; Calliphora.  
OX NCBI\_TaxID=7373;  
[1]  
RN SEQUENCE FROM N.A., AND CHARACTERIZATION.  
RP STRAIN=CHALKY; TISSUE=RETINA.  
RX MEDLINE=96216082; PubMed=8662634;  
RA Huber A., Sander P., Paulsen R.;  
RT "Phosphorylation of the Inad gene product, a photoreceptor membrane protein required for recovery of visual excitation".  
RL J. Biol. Chem. 271:11710-11717(1996).  
CC -FUNCTION: MAY BE INVOLVED IN CONTROL OF THE LIGHT RESPONSE. A RISE  
IN INTRACELLULAR CALCIUM LEVELS UPON VISUAL EXCITATION MAY  
INITIATE PHOSPHORYLATION OF THE INAD PROTEIN BY EYE-PKC.

CC PHOSPHORYLATED INAD MAY IN TURN ACT ON ANOTHER PROTEIN SUCH AS TRP  
 CC OR NORIA WHICH BOTH CO-PRECIPIRATE WITH INAD.  
 CC -1- SUBCELLULAR LOCATION: MEMBRANE-ASSOCIATED.  
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN RHADOMERAL PHOTORECEPTORS.  
 CC -1- Ptm: PHOSPHORYLATED BY EYE-PKC (PROBABLE).  
 CC -1- SIMILARITY: TO DROSOPHILA INAD.  
 DR EMBL: 269889; CAA93758.1; .  
 DR HSSP: P31016; 1BFE.  
 DR InterPro: IPR001478; .  
 DR Pfam: PF00595; PDZ; 5.  
 DR SMART; SM00228; PDZ; 1.  
 DR Repeat; Phosphorylation; Glycoprotein.  
 FT DOMAIN 41 420 2 X APPROXIMATE REPEATS.  
 FT REPEAT 41 82 1.  
 FT CARBOHYD 381 420 2.  
 FT CARBOHYD 5 5 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 162 162 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 272 272 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 292 292 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 303 303 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT MOD\_RES 15 15 BY PKC (POTENTIAL).  
 FT MOD\_RES 168 168 BY PKC (POTENTIAL).  
 FT MOD\_RES 185 185 BY PKC (POTENTIAL).  
 FT MOD\_RES 320 320 BY PKC (POTENTIAL).  
 FT MOD\_RES 321 321 BY PKC (POTENTIAL).  
 FT MOD\_RES 437 437 BY PKC (POTENTIAL).  
 FT MOD\_RES 595 595 BY PKC (POTENTIAL).  
 FT MOD\_RES 658 658 BY PKC (POTENTIAL).  
 SQ SEQUENCE 665 AA; 73349 MW; C453192E18202BF7 CRC64;

Query Match 8.1%; Score 564.5; DB 5; Length 665;  
 Best Local Similarity 26.8%; Pred. No. 1.7e-25;

Matches 190; Conservative 109; Mismatches 229; Indels 181; Gaps 27;

OY 521 OPRARVLEMPKSLGSLVIGRGKSGRLSNGEVRGIFIKHLEDSPAKNGTLKPGDR 580  
 DB 10 QVQSVTLDRKTKKSFGLSLVRGEA-----RDGSKNGIFIKGIIVPSPGLGKIVGDR 64  
 OY 581 IYBAPS---QSESEPE----- 593  
 DB 65 LTLNKGVDVDAPEVILNIKQAGSKIDLEOTYGEQSNKDKNGEIKENGESQNNMEN 124  
 OY 594 ----KAPLCVSP-----PPPSAFEMGSDHTOSSAS----- 621  
 DB 125 QDSIKOPTKQPAIKQOSKQAPGRPPVNNAMLNKSNFTTSSKSDONLDDDEPTDRM 184  
 OY 622 -----KISQDYDK--EDEFYGMKNIRERYGLTGLHMIEL 656  
 DB 185 TGRIRTAAGEIDRASAGNCKLNKMEKDKETDEFGYTKAKINKRYTMR-DLKKLEI 243  
 OY 657 EK-GHSGLSLAGNCKRSMVSFYIVGIDPNC-AAGKDLQIADLELNGOILYGRSH 714  
 DB 244 VRTNTALILALAGHSDRCKMCFVAGVNTSGFLASVD--IKSGDEILEVNGVTLKNRCH 301  
 OY 715 QNASSTIKCAPSKVILIFIRNKDAVNQMAVCPGNAVEPLPSNSENLOKNETEPTVTSDA 774  
 DB 302 LNASVIFKINDE-KLVLTISRKKPND-----EGMSVKKP-----KKFPEI--DDT 345  
 OY 775 AVDLSSFKNVQHLKPKDGGGLGIAI---SEEDTLGVIKSLTEHGAATDRLKLVGO 831  
 DB 346 KLFEGEYAKARASYV-KKGGFLGIMVYKQHVNGGIFISDLRESNMLAG-LAVGDM 403  
 OY 832 ILAVALDEIYVGVIEKIFISLTKAKMTVKLTTHAENPDQOANPSAAGAAGEKKNSQSL 891  
 DB 404 LLAIVNDVCEVSNYDAVALLKRAEGIVNLVLTLETDI-----KKEEEE 452  
 OY 892 WYPOSGP-EPEIRNTSSRSFPAIPASDPATCPIIPGCEITIEISKRTGLSLTVGS 950  
 DB 453 KKKEAARPEPEK-----PVDSIIEVPGRIYEVATDKKPLGIVYVGK 499  
 OY 951 DTLL--GATIEHYEYEGAAKDKRLMAGDQILEVNGIDLRKATHDIAINVL----- 1001

DB 500 NNVTGCTIITH-IYEGVIAEDKRLKIPDHIIQVNG----KEVQCEANTLKVHQLFYT 554  
 OY 1002 QTPQARRLLYDEAPYKKEEVCOTITITLQKPKGGLSLYGRNDGVSVDIWKG 1061  
 DB 555 LYEKVITIVYRADPP---EV-ETFKVEFARKAGDGLSLA--PNEGCTISSETTSAG 607  
 OY 1062 IADPDRLIQDQIILLVNGEDVRNMAQEAVALKCSLSTVLEVGRIK 1110  
 DB 608 YADIDNKLORGDIITKFNQDSLEGLTFEYCALFGATGATKISLEITRPK 656

RESULT 14  
 O9NBV3 PRELIMINARY; PRT; 674 AA.

AC O9NBV3  
 DT 01-OCT-2000 (TREMBLrel. 15, Created)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
 DE 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
 GN INAD.  
 OS INAD OR CG3504.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RC SEQUENCE FROM N.A.  
 RP STRAIN=CANTON-S;  
 RA Ashari M., Shieh B.;  
 RT "The Inad locus in Drosophila."  
 RL Submitted (MAR-2000) to the EMBL/Genbank/DBD databases.  
 DR EMBL: AF245280; AAF81203.1; .  
 DR Flybase; FBgn001263; Inad.  
 DR InterPro; IPR00582; .  
 DR InterPro; IPR001478; .  
 DR Pfam; PF00595; PDZ; 5.  
 DR Prodom; PD002965; .  
 DR SMART; SM00228; PDZ; 1.  
 SQ SEQUENCE 674 AA; 74305 MW; 55B1C57B7A100FID CRC64;

Query Match 8.0%; Score 556; DB 5; Length 674;  
 Best Local Similarity 26.5%; Pred. No. 5.7e-25;

Matches 202; Conservative 134; Mismatches 292; Indels 134; Gaps 31;

OY 368 VMLQNVSKSEFERTINIAKNSLSLGTAVSANKDGLMIVRSIIHGAIISRDRIATGCI 427  
 DB 18 VTLDKTKKSF--GLCIVRGE---VKDSPNTKTTGIFIKGIIVPSPAHLCRLKVGDR 71  
 OY 428 LSLNVESTSVNNAQARMLRHRSLILGPDIKITTYAHELEFKISLQOOSGRVMAIDF 487  
 DB 72 LSLNGK---DVNRSTEQAVI--DLIKE-----ADFKIEL-----ELQTFD-- 106  
 OY 488 SSYTGRIPELPEREGESEELQNTAYSNNOPRVELMREPSKSLGSIYGRGMS 547  
 DB 107 -----KQDEQAKSDPRSNGYMAKKNFNDQOTNNAS--GCGGMO 147  
 OY 548 RLSNGEVRMGIFIKHLEDSPAKNGTLKPGDRIYAPQSQSESEPEKAPLCVPPPPSA 607  
 DB 148 GOGGOGGAGM-----NRQOSMOKRNTTFTASMKQKSHNADEDEDEPTRMTCG----RI 197  
 OY 608 PLEMGSDHROSSA-----KISQDYDK--EDEFYGMKNIRERYGLTGLHMIELKXH 660  
 DB 198 RTBAGYIDIRASAGNCKLNKOEKDKEDDEFGYTKAKINKRYTMR-DLRIEYQDA 256  
 OY 661 S-GIGSLAGNCKRSMVSFYIVGIDPNGAAGKDLQIADLELNGOILYGRSHONASS 719  
 DB 257 SKPLGLALAGHDKRQMACFVAGVDPNGALGVD--IKPDDELIVGVNLKRNCHLMA 315  
 OY 720 IIKCAPSKVILIFIRNKDAVNQMAVCPGNAVEPLPSNSENLOKNETEPTVTSDA 779  
 DB 316 VRSVVDG--KLWMTISRKRPNDEGMC---VRFI-----KKFPASDETKFI--FD 359

QY 780 SFKNVHLELPKDOGLGIAI--SEEDTSGVIKSLTEHGVAATDRLKVGDOILAVD 836  
 Db 360 QPKRATVQVRK-EGFGLIMYIVYKHAIEVSGIFISDLREGSNMELAG-VKVGMDLVLAV 417  
 QY 837 DEIVGVYIEKFIISLTKAKTKV--LTHAENDSNAVSAAGAGKKNSSQSLMV 893  
 Db 418 QDVLESYWDATGTLKAEVYVMIITLTKSE---EAIKAEKAEKKEEAKKEEK 473  
 QY 894 POSGSEPESEIRNTSRSTPAIFASDPATCPILPCETTEISKRTGLSIVGSD-- 951  
 Db 474 PQ-----EPATAEIKPKKLLILEKVKRMGYIVCGCKNNH 510  
 QY 952 TLLGAFITHEVYEGSACKDRLNAGDOILEVNG--IDLRAITHDEALNLRQIPQR-VR 1008  
 Db 511 VTTGCVITH-VYPEGQVADRLKFKFDHICDINGTPHVGSMTELLKVLQFLPHTYEKAVT 569  
 QY 1009 LTVIRDEAPYKEEYECDLTLELOKKPKGGLSIVGKRNTOGVPSIVGSIADPDGR 1068  
 Db 570 LTVFADPPELEK-----FVNDLKKKAKKELGLSL--SPNEIGCTIADLIGQYPEIDSK 622  
 QY 1069 LIQGDQILLVNGEDVNRASQEAVALLKCSLGTVLEYGRK 1110  
 Db 623 LQRGIIITKFPNGDALEGLPFOVCYALFKGANGKSMETTRK 664

RESULT 15  
 Q24008 PRELIMINARY; PRT; 674 AA.  
 AC Q24008  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)  
 DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)  
 DE INAD PROTEIN.  
 GN INAD OR CG3504.  
 OS Eukaryota: Metazoa: Arthropoda: Tracheata: Hexapoda: Insecta:  
 OC Pterygota: Neoptera: Endopterygota: Diptera: Brachycera: Muscomorpha;  
 CC Ephydriidae: Drosophilidae: Drosophila.  
 CX NCBI TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BERKELEY:  
 RA MEDLINE-20196006; PubMed-10731132;  
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazek R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt J., Nelson C.R., Miklos G.L.G.,  
 RA Abell J.F., Agbayani A., An H.-J., Andrews-Piankocch C., Baldwin D.,  
 RA Baller R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Bokora D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,  
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cavley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pallos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostalka D., Houston K.A., Howland T.J., Wei M.-H., Idegawa C.,  
 RA Jostali M., Kaulush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Mekulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pachle J.M.,  
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheffler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,

RA Svrlkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zhang X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of Drosophila melanogaster";  
 RL Science 287:2185-2195(2000).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-COMPOUND EYE.  
 RA Lindsley D.L., Zimm G.G.;  
 RT "The genome of Drosophila melanogaster";  
 RL submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-COMPOUND EYE.  
 RA MEDLINE-95127229; PubMed-7826638;  
 RA Shieh B.H., Niemeyer B.;  
 RT "A novel protein encoded by the Inad gene regulates recovery of visual  
 RL transduction in Drosophila.";  
 RL Neuron 14:201-210(1995).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-COMPOUND EYE.  
 RA Shieh B.;  
 RL submitted (OCT-1994) to the EMBL/GenBank/DBJ databases.  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-COMPOUND EYE.  
 RA Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.  
 RL EMBL: AE003458; AAF46915.1;  
 DR EMBL: U15803; AAC36490.1;  
 DR HSSP: P31016; 1BFE.  
 DR FlyBase: FBgn0001263; Inad.  
 DR InterPro: IPR001478;  
 DR Pfam: PF00595; PDZ; 5.  
 DR SMART: SM00228; PDZ; 1.  
 SQ SEQUENCE 674 AA; 74332 MW; DAC24091D9EA7F3 CRC64;

Query Match 7.98; Score 554; DB 5; Length 674;  
 Best Local Similarity 26.58; Pred. No. 7.3e-25;  
 Matches 202; Conservative 134; Mismatches 292; Indels 134; Gaps 31;

QY 368 VMLQNSKESPERINTAKGSSLGMTVSANKDGLGMYRSHIHGALSRGRIADICI 427  
 Db 18 VTLKRTGKSF--GICIVRGE-----VKDSPYTKTGTIFIKGIVDPSPAHLCGRKVGDR 71  
 QY 428 LSINESTISVTNQAARMLRRHSLIGPDIKITYVPAHELEFKISLGQSGRYMALDIF 487  
 Db 72 LSLNGK---DVNRNTEQAVI---DLIKE-----ADRIEL-----ETQTFD-- 106  
 QY 488 SSYGRDIPELPEREEGEGESELQNTAVSNMOPRRVYELWRPESKLSIGISVGRMG 547  
 Db 107 -----KSDQAKSPDRSNGYQAKKFNQEOGTNNNAS--GGQGMQ 147  
 QY 548 RLSNGEVRGFIYHVLDEDSAGKNGILKPRDRIYEAPSQSESEPEKAPLCISVPPPPSA 607  
 Db 148 GQGGQGMAGN---NRQSQMKRNTTFTASMRKHSYVADDEDTRDMG-----RI 197  
 QY 608 FAEGSGHSTOSSA-----SKTSDPVR--EDEFYSKKNIRERYGTLGELHMLEKGH 660  
 Db 198 RTEAGYEIDRASAGCNKLINQERKDEDEFEYTAKKIRKNNMK--DLARILEYORDA 256  
 QY 661 S-GGLSLAGKDRSRMSVFLVIGDPNGAAGKDRLOIADLEELINGQILYGRSHQNAS 719  
 Db 257 SKPLGLALAGHKDRQAKAFVAGVDPNGALGSVD--IKRGEIYVGNVNLKRNCHLNASA 315  
 QY 720 IIKAPSKVITIFIRNNDVAVQMAVCGNAVEPLPSNSENLQNKETEPYTTSSAAVDLS 779  
 Db 316 VFKNVGD-KIVMTISRKPDEGMC---VKPI-----KKPPTASDETTFI--FD 359

```

OY 780 SFKNVQHLELPKDGGLGIAI---SEEDTSGVITIKSLTEHGVAAATDGRKVGDOILAVD 836
Db 360 QPFRKFTVQVRK-EGFLGIMVIYGHAEVSGSIFISDLREGSNAELAG-VKVGDMFLAVN 417
OY 837 DEIVGVPIEKFTISLTKTAKMTVK---LTIHAENPDQAVPSAGAGASGEEKNSQSILMV 893
Db 418 QDVLTESNVDATGTLKRAEGVYTMILTLKSE---EAIKAKAAEKKKKEAKKEEK 473
OY 894 POSGSPESPESIRNTSRSSSTPALFASDPATCPIIPGCEFTTIEISKGRGTGLSLIVGSD-- 951
Db 474 PQ-----EPATAETKPNKILIELKVEKPKPMGVIVCGKNMH 510
OY 952 TILGAFIHEVYEBGACDGRUMAGDOILEVNG--IDLKRAITHDEAINVLRQTPQR-VR 1008
Db 511 VTTGCVITH-VYPEGOVADKRLKIFDHICDINGTPIHVGSMITLKVHQLFHTYEKAVT 569
OY 1009 LTLVHDEAPYKEEVEYCDTLTIELQKPKGKGLSIVGKRNDTGVFVSDIVKGIADPDGR 1068
Db 570 LTVFRADPELEK-----FNVDLMKKAKGKELGLSL--SPNEIGCTIADLIQGYPEIDSK 622
OY 1069 LIQSDQILLVNGEDVYRNASQEAVALLKCSLGTVTLEVGRIK 1110
Db 623 LQRCDIITKFNGBDALEGLPFOVCYALFKGANGKVSMEVTRPK 664

```

Search completed: July 12, 2001, 14:43:01  
 Job time: 214 sec

•  
•  
•  
•

W

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2001, 14:37:37 ; Search time 23.36 seconds

(without alignments)  
1184.015 Million cell updates/sec

Title: US-09-502-698-1

Perfect score: 6990

Sequence: 1 MWCRRVPTTQSELDLD.....EAAVALIKRTKGYTLNVL 1373

Scoring table:

Gapop 10.0 , Gapext 0.5

Searched: 193259 seqs, 20144635 residues

Total number of hits satisfying chosen parameters: 193259

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

Issued\_Patents\_AA:\*  
1: /cgn2\_6/ptodata/2/1aa/5A.COMB.pep:\*  
2: /cgn2\_6/ptodata/2/1aa/5B.COMB.pep:\*  
3: /cgn2\_6/ptodata/2/1aa/6A.COMB.pep:\*  
4: /cgn2\_6/ptodata/2/1aa/6B.COMB.pep:\*  
5: /cgn2\_6/ptodata/2/1aa/PTUS.COMB.pep:\*  
6: /cgn2\_6/ptodata/2/1aa/Dackfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | ID               | Description       |
|------------|-------|-------------|--------|------------------|-------------------|
| 1          | 509   | 7.3         | 2485   | US-09-290-640-46 | Sequence 46, Appl |
| 2          | 500.5 | 7.2         | 2466   | US-09-080-855-12 | Sequence 12, Appl |
| 3          | 500.5 | 7.2         | 2466   | PCT-US94-09943-2 | Sequence 2, Appl  |
| 4          | 498   | 7.1         | 2465   | US-08-596-291-3  | Sequence 3, Appl  |
| 5          | 498   | 7.1         | 2465   | US-09-100-804-3  | Sequence 3, Appl  |
| 6          | 340   | 4.9         | 1112   | US-09-045-632-2  | Sequence 2, Appl  |
| 7          | 340   | 4.9         | 1112   | US-09-045-632-3  | Sequence 3, Appl  |
| 8          | 335   | 4.8         | 1018   | US-09-045-632-15 | Sequence 15, Appl |
| 9          | 335   | 4.8         | 1018   | US-09-045-632-16 | Sequence 16, Appl |
| 10         | 335   | 4.8         | 1061   | US-09-045-632-32 | Sequence 32, Appl |
| 11         | 334.5 | 4.8         | 610    | US-08-410-804-1  | Sequence 1, Appl  |
| 12         | 334.5 | 4.8         | 610    | US-08-259-514-1  | Sequence 1, Appl  |
| 13         | 334.5 | 4.8         | 610    | US-08-858-311-1  | Sequence 1, Appl  |
| 14         | 334.5 | 4.8         | 1050   | US-09-045-632-50 | Sequence 50, Appl |
| 15         | 331.5 | 4.7         | 1050   | US-09-045-632-49 | Sequence 49, Appl |
| 16         | 329.5 | 4.7         | 604    | US-09-045-632-14 | Sequence 14, Appl |
| 17         | 327   | 4.7         | 918    | US-09-045-632-21 | Sequence 21, Appl |
| 18         | 327   | 4.7         | 961    | US-09-045-632-33 | Sequence 33, Appl |
| 19         | 321   | 4.6         | 602    | US-09-045-632-20 | Sequence 20, Appl |
| 20         | 288.5 | 4.1         | 818    | US-09-045-632-25 | Sequence 25, Appl |
| 21         | 287   | 4.1         | 818    | US-09-045-632-13 | Sequence 13, Appl |
| 22         | 273   | 3.9         | 504    | US-09-045-632-19 | Sequence 19, Appl |
| 23         | 271   | 3.9         | 504    | US-09-045-632-19 | Sequence 19, Appl |
| 24         | 266.5 | 3.8         | 642    | US-09-045-632-35 | Sequence 35, Appl |
| 25         | 265   | 3.8         | 599    | US-09-045-632-28 | Sequence 28, Appl |
| 26         | 258.5 | 3.7         | 502    | US-09-045-632-24 | Sequence 24, Appl |
| 27         | 252   | 3.6         | 374    | US-09-091-405-2  | Sequence 2, Appl  |

|    |       |     |      |                   |                   |
|----|-------|-----|------|-------------------|-------------------|
| 28 | 227   | 3.2 | 519  | US-08-997-445D-2  | Sequence 2, Appl  |
| 29 | 225.5 | 3.2 | 298  | US-09-045-632-27  | Sequence 27, Appl |
| 30 | 219   | 3.1 | 498  | US-09-045-632-30  | Sequence 30, Appl |
| 31 | 219   | 3.1 | 541  | US-09-045-632-36  | Sequence 36, Appl |
| 32 | 217   | 3.1 | 284  | US-09-045-632-12  | Sequence 12, Appl |
| 33 | 214.5 | 3.1 | 407  | US-09-045-632-18  | Sequence 18, Appl |
| 34 | 209.5 | 3.0 | 1612 | US-08-545-860D-48 | Sequence 48, Appl |
| 35 | 209.5 | 3.0 | 1612 | PCT-US94-04496-48 | Sequence 48, Appl |
| 36 | 203   | 2.9 | 404  | US-09-045-632-23  | Sequence 23, Appl |
| 37 | 200   | 2.9 | 79   | US-09-100-804-27  | Sequence 27, Appl |
| 38 | 200   | 2.9 | 86   | US-08-545-860D-53 | Sequence 53, Appl |
| 39 | 196.5 | 2.9 | 86   | PCT-US94-04496-53 | Sequence 53, Appl |
| 40 | 196.5 | 2.8 | 198  | US-09-045-632-11  | Sequence 11, Appl |
| 41 | 190.5 | 2.7 | 1829 | US-09-157-420-11  | Sequence 1, Appl  |
| 42 | 178.5 | 2.6 | 182  | US-09-045-632-29  | Sequence 29, Appl |
| 43 | 176.5 | 2.5 | 233  | US-09-151-611-1   | Sequence 1, Appl  |
| 44 | 173   | 2.5 | 505  | US-08-123-161A-14 | Sequence 14, Appl |
| 45 | 173   | 2.5 | 505  | US-08-483-278-14  | Sequence 14, Appl |

#### ALIGNMENTS

RESULT 1  
US-09-290-640-46  
Sequence 46, Application US/09290640  
Patent No. 6204055

GENERAL INFORMATION:

APPLICANT: Dean, Nicholas M.

FILE OF INVENTION: Marcussen, Eric G.

FILE REFERENCE: ISPH-0351

CURRENT APPLICATION NUMBER: US/09/290,640

CURRENT FILING DATE: 1999-04-12

NUMBER OF SEQ ID NOS: 85

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 46

LENGTH: 2485

TYPE: PRT

ORGANISM: Homo sapiens

US-09-290-640-46

Query Match 7.3%; Score 509; DB 4; Length 2485;  
Best Local Similarity 22.8%; Pred. No. 36-32;  
Matches 243; Conservative 140; Mismatches 371; Indels 314; Gaps 37;

|    |      |  |      |
|----|------|--|------|
| QY | 530  | EPKSLGISIVGGRGMSRLNCEVNRGIFIKHLEDPAGKNGTLKPGDRIVEAPSOSE  | 589  |
| DB | 990  | EPPTVAELVGRKPSHMSRSDAELAGV---TKLNKSVASLNRSPERRKHSSSSSI   | 1046 |
| QY | 590  | SEPKAPLCSVPPPPSAFEMGSDHTQSSAKISQDVDEDFGYSWKNIRRYGTILG    | 649  |
| DB | 1047 | EDPGAGVYVIGMT-----MHSSGNSSOVPLKENDV-----LHRMWSIVS        | 1086 |
| QY | 650  | ---ELMIIELEK-GHSGIGLSLGNKDRSR--NSVFTVIGIDPGANGKGRQIADELE | 703  |
| DB | 1087 | PEREITLVNKKAKYGLGQITIGGEMGRLDGITISSVAPGPADLDGCLPGRLIS    | 1146 |
| QY | 704  | INQILYGRSHQWASSIIKCAPSKVRIIFRNDAVQNAVCNANVEPLPSNENLQNK   | 763  |
| DB | 1147 | VNSVSLGYSVSHNAIETLQNAPEDEVLIQSPEKISKVSTVH---LTNEMKRYMVK  | 1202 |
| QY | 764  | ETEPYVTSDAVDSLSPFN-----VOHLEPK--DQGL--GIAISEEDTLGVIYS    | 813  |
| DB | 1203 | SS---YMODSAIDSSKHDMGRGLRHISENSFPSPGLRGSLSDSPRESASLSQ     | 1258 |
| QY | 814  | LTEHGVAATDGRKLVQDQILAVDEIVYGVPIKFTSLKTAQVYKLTIAENPSONAV  | 873  |
| DB | 1259 | SOVNGFPAS---HLDQVW---QESOHGSPSPSVIS---KATEKEFTDSNQSKTK   | 1306 |
| QY | 874  | PSAAGAASGEKK-----NSSOSLWPGQSGPESPEIRTSRSRSPAIRPASPPATC   | 923  |

Db 1307 PGISDVPTYSRGRGSDMDDEATYSSSDQHQPWKQ---ESSSSVNTSNKMNKFTSSSPK- 1362

Qy 924 PIIRGCETTIITSKGRGGLGSLTYGSSDTLL--GAFIIHEYEGAACTKGRIMAGQIL 961

Db 1363 ---PGDIEFVELANDNSLISVTGVTNVRHGIVYKAVIPIOGAAESDGRHKGRVLL 1419

Qy 982 EVNGIDLRKARHDEAIINVLROTPOVRRLTYIRDAPYKEEV---C----- 1024

Db 1420 AVNGVSLGKATHKQAVELRLGTGVHLLLEKSGPSKHEIVPTPCTISDQNAQGGP 1479

Qy 1025 -----DVLTELOKPKGKGLS-----IVGRNDTGVFSDIVK 1059

Db 1480 EKVAKTTQVKDYSFVTEENETFEVKLFKN-SGGIGFSFSREDNLLPEQINASIYRVKKLFP 1538

Qy 1060 GGIDPDSRLQGGQOILLVNGEDVRNNAQOEAVALLCSGLTYLLEGRIRAK----- 1112

Db 1539 GQPAAESKIDVGVDTILKVNASLKLGLSQDEIVSALRGTAPEVFLLCRPPGLPELDT 1598

Qy 1113 ---PFHSE-----RRPSQTSQVSEGS-- 1130

Db 1599 ALTRPLQSPAQVLPNNSKDSQPCVBPQSTSDENEMSKDKKQCKSPSRDSYSDSSGS 1658

Qy 1131 ---LSSFTF-----PLSGST----- 1143

Db 1659 GEDDLVTAPANISNSTWSALHQTLSNMVSOAOSHHBAPRSQEDTCTMFEYPOKIPNKP 1718

Qy 1144 -----SESSLESSK-----KNALASEI 1160

Db 1719 EFEDSNPSPLRPDMAPOGQVQPOSESASSSMCKYTHIHISEPTROENMTPLKNDLENHL 1778

Qy 1161 QGLR-----IYEMKKGPPLDSLGISIAG--VGSPLCDVPIFLAMHPTGVAATOGLRV 1212

Db 1779 EDFEVELLTLTKSEKSGLGFVTGKGNRIGGYADV---IDPP--AASDGLRK 1830

Qy 1213 GDRIVTYIGTSTEGMTHTQAVNLKKNASGSIEMQVAVAGDVSVYTGHOE-PASSLSFT 1271

Db 1831 GDRILKVNDDTYTMTHTDAVNLRAASKYRL-----YIGVLELPIPIML--- 1877

Qy 1272 GLTSTSTIFODDLGPCKSITLERGPDGLGFSFISVGYGSPHGLPIYKTVFAKGAASD 1331

Db 1878 -----PHLLPDITLTCNKEELGFSLCGGHDSLY--QVYYSIDINPRSYAIE 1922

Qy 1332 GRLKRGQIIIVANGOSLEGVTHEA-----VALTKRTKGVITLM 1370

Db 1923 GNLDLLDVIIHYVNGVSTOGMTLEEVRNALDMSLPVLVIAKATRDNDLPV 1970

RESULT 2

US-09-080-855-12

Sequence 12. Application US/09080855A

Patent No. 6083721

GENERAL INFORMATION:

APPLICANT: Saras, Jan

APPLICANT: Franz, Petra

APPLICANT: Aspenstrm, Pontus

APPLICANT: Hellman, Ulf

APPLICANT: Genez, Leonel Jorge

APPLICANT: Heidlin, Carl-Henrik

TITLE OF INVENTION: PARC, A GPBASE ACTIVATING PROTEIN WHICH INTERACTS WITH PTPLP

FILE REFERENCE: 10461/7030

CURRENT APPLICATION NUMBER: US/09/080,855A

CURRENT FILING DATE: 1998-05-18

EARLIER APPLICATION NUMBER: 08/805,583

EARLIER FILING DATE: 1997-02-25

NUMBER OF SEQ ID NOS: 39

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 12

LENGTH: 2466

TYPE: PRT

ORGANISM: Homo sapiens

US-09-080-855-12

|    | Query Match | Similarity           | 7.2%                         | Score 500.5:       | DB: 3               | Length 2466;   |                  |                |           |
|----|-------------|----------------------|------------------------------|--------------------|---------------------|----------------|------------------|----------------|-----------|
|    | Best Local  | Similarity           | 22.2%                        | Pred. No. 1.5e-31; |                     |                |                  |                |           |
|    | Matches     | 253;                 | Conservative                 | 146;               | Mismatches          | 396;           |                  |                |           |
|    |             |                      |                              |                    | Indels              | 351;           |                  |                |           |
|    |             |                      |                              |                    | Gaps                | 41.            |                  |                |           |
| QY | 510         | ELONAYSNMNO          | PRRVELMREPSKLGISYIGRMGSR     | LSNG-----          | 552                 |                |                  |                |           |
| Db | 868         | QLOMARQSNDAQO        | IERASFRLNLQAESVGFMMGRALISG   | LASSPLNKLA         | PLSLV               | 927            |                  |                |           |
| QY | 553         | --EVRKGFILKHV----    | LEDSPACKNGL----              | KPGD--             | RIVEARPSQESPEEAPLCS | 599            |                  |                |           |
| Db | 928         | QAEILKRLSCSEL        | SLYOPLONSKREKRNKASWEKEPR     | REMSKYHDL          | SOASLYPHKRNIV       | 987            |                  |                |           |
| QY | 600         | VPPPSAFEAEM          | -GSDHTQSSAS-----             |                    |                     |                |                  |                |           |
| Db | 988         | NMEPPQVVAL           | ELYGRPSHQMSRSDAESLAGYTKLNKNS | SVASLNSP           | PERKHESSSSSIE       | 1047           |                  |                |           |
| QY | 632         | EFG--YSKNIRERYGTLG-- | --ELHMLELEK--                | GHSGLGLSLAGK       | DNRSR--MSVFIYGI     | 683            |                  |                |           |
| Db | 1048        | DPGQAVLDV            | LHKRMSIYSSPEREITLVNKKDKAY    | GGLGFQIITGEEKMRDL  | IGLIGFISV           | 1107           |                  |                |           |
| QY | 684         | DPNGAAGDGR           | QIODELLEINQGLLYGSHONASSIIR   | CAASKYKII          | FIRNKDAVNOMA        | 743            |                  |                |           |
| Db | 1108        | ARGGADHRC            | LKRPDRILSVASVSEGVSHHALEILL   | ONAPEDYTLV         | ISQKEKISKVP         | 1167           |                  |                |           |
| QY | 744         | VCPGAVER             | ELPDSNENLONKETEPTVTTSDAAV    | DLSSFRN-----       |                     |                |                  |                |           |
| Db | 1168        | STPVM-----           | LTNEMKNMYKKS-----            | WQDSALIDSSSDHMSR   | GTLLHISNESP         | PGSG 1219      |                  |                |           |
| QY | 796         | L-GIAISEED           | PLSGVILKSLTEHGVANDGRLKYGDO   | ILAVDELIVG         | YPERKFI             | SLK 853        |                  |                |           |
| Db | 1220        | LREGLSODS            | DKTEBASLSQSOVNFPS-----       | HLGDTW--           | -QDSQHSPE           | PVSIS--- 1269  |                  |                |           |
| QY | 854         | TAKMTVKLT            | LHAENDSQAVPSAAGAAGEK-----    |                    |                     |                |                  |                |           |
| Db | 1270        | --KATEKETFD          | DSNOSKTKKPGISDVTYSDRGSD      | MDMEATYSSQD        | QHPKXO--            | -----ESSS 1324 |                  |                |           |
| QY | 904         | IRNRSRST             | PAIFADPANCPIPGCEITIEKSR      | TGIGLSIVG          | SOTLL--GAFI         | IHE 961        |                  |                |           |
| Db | 1335        | SVNYSNKN             | NFTFESSPR-----               | PDITFEVLAK         | DNSSLGISV           | TGCVNTSVHHG    | ITVKA 1380       |                |           |
| QY | 962         | VEEGSAACK            | DGLWAGDOILEVNGIDLEK          | KATHEAIVNLROT      | POKRVRLTY           | REAPYKEE 1021  |                  |                |           |
| Db | 1381        | VIPQGAAS             | SDGRIRHGDVYLANGVSLBG         | ATKQAVETLR         | NTGCVVHLL           | LEKGGSP        | TKE 1440         |                |           |
| QY | 1022        | EY-----C-----        |                              |                    |                     |                |                  |                |           |
| Db | 1441        | HVPVTPQ              | CTLSDQNAAGQPEKVKTKT          | YQKDYSFYTE         | ENTEFEV             | KLFKN--SSGLG   | SIFGRE 1499      |                |           |
| QY | 1043        | ---IYGRND            | GVFESDLYVKGGIADPDR           | LIOGDQILLVNG       | EDVYNASQEA          | VAALLKCSL 1099 |                  |                |           |
| Db | 1500        | DNILPEQ              | INASTIVRKRLFACQPA            | AESGRKIDVGYIL      | KVNGASL             | GLSGOQEV       | YISALRGA 1559    |                |           |
| QY | 1100        | GTVLEVR              | KIKAG-----                   | PFHSE-----         |                     |                |                  |                |           |
| Db | 1560        | PEVFLIL              | CRPPGVLPETIDTAL              | PLPQSPAOVLP        | NSKSDSQSP           | CEVOSTS        | DEMEMSDK 1619    |                |           |
| QY | 1118        | -----RRSQ            | SQSVSEGS-----                |                    |                     |                |                  |                |           |
| Db | 1620        | SKKQCKS              | PSRKRDSYSSGSGED              | LVTPANIS           | NSTWSAL             | HQTLISNMV      | SOAQSHHEAPK 1679 |                |           |
| QY | 1139        | SGSST-               |                              |                    |                     |                |                  |                |           |
| Db | 1680        | SOEDICT              | MFYYPQKIPRKPEFED             | NSPPLPDPMA         | PGQSYOPQ            | SSBASSSMD      | KYIHII 1739      |                |           |
| QY | 1153        | -----KALAE           | IEIGLR-----                  | YEMKKGPR           | DSLGIS              | TAGC--VGS      | PLGVP 1192       |                |           |
| Db | 1740        | SEPTROEN             | WPLKNDLEHLED                 | FELEVELL           | TLTLKSE             | KASLGFTY       | TKNQRIGCV        | YHV- 1798      |           |
| QY | 1193        | IFIAMHPT             | GTAAQTKRLVGDRI               | YITIGSTEG          | HTQAVLLK            | KNAGSIE        | MQVAVAGD 1252    |                |           |
| Db | 1799        | ----IQD--            | -----AKSDGR                  | LKRPDRIL           | KVNDIVT             | YNTNTH         | TDANVLLRAAK      | YVRL----- 1844 |           |
| QY | 1253        | VSVYTG               | HHQOE--PASSSL                | SFTGLTSTIS         | FODDLG              | PPQCKS         | ITLERKPGDLG      | FSIVG          | YGSP 1311 |



Db 1845 ---VIGRVLPRIPML-----PHLLPITLTCKNEELGFSICGHDLSL 1885  
Qy 1312 HDDLPIYKTVAKGASDGLKRGDQILIANVNGSLBEVTHEEA-----VAILKR 1362  
Db 1886 Y--QVYVSDIPRSVAALIEGNLQDLIDVHYHNGVSTQGMTEEVNRALDMSLPSLVLKA 1943  
Qy 1363 TKGVTLM 1370  
Db 1944 TRNDLPV 1951

RESULT 3  
PCT-US94-09943-2Sequence 2, Application PC/TUS9409943  
GENERAL INFORMATION:

APPLICANT:

APPLICANT:

APPLICANT:

APPLICANT:

TITLE OF INVENTION: PRIMARY STRUCTURE AND FUNCTIONAL

TITLE OF INVENTION: EXPRESSION OF NUCLEOTIDE SEQUENCES FOR NOVEL PROTEIN

NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:

CITY: BOSTON

STATE: MASSACHUSETTS

COUNTRY: USA

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US94/09943

FILING DATE: 01-SEP-1994

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/115,573

FILING DATE: 01-SEP-1993

ATTORNEY/AGENT INFORMATION:

NAME: TWOMEY, MICHAEL J.

REGISTRATION NUMBER: P-38,349

REFERENCE/DOCKET NUMBER: L0461/7000WO

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617/720-3500

TELEFAX: 617/720-2441

TELEX: 92-1742 EZEKIEL

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 2466 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

PCT-US94-09943-2

Query Match 7.2%; Score 500.5; DB 5; Length 2466;  
Best Local Similarity 22.2%; Pred. No. 1.5e-31;  
Matches 255; Conservative 146; Mismatches 396; Indels 351; Gaps 41;

Qy 510 ELQNTAISMNMPRVELMREPSKSLGISVGRGMGKSLNG----- 552  
Db 868 QLOMAROSNODADODIERASFRSLNLQASVYRGFMNGRAISTGSLASTLNLKLAVRPLSV 927  
Qy 553 --EVARGIFIKHV-----LEDSPAGKNGTL--KPGD--RIVEAPSOSESEPEKAPLCS 599  
Db 928 QAEILKRLSCSLSLXQPLQNSKREKNDKASWEKPREMSKYHLSQASLTPHKKNVY 987  
Qy 600 VPPPPSAFAEN-GSDHTQSSAS-----KISQDVKED 631

Db 988 NMEPPQVAVELVCKPSHOMSRSDAELAGVTKLNNKSVASLNSPERKHHESDSSIE 1047  
Qy 632 ERG--YSKNNIERGTLTG---ELHMLEK-GHSGGLSLAGKNDKR--MSFIVGI 683  
Db 1048 DPGQAVLDLHKLKMSYSSPEREITLVNLKRDARYGLAFOIIGCKMRDLGIFISSV 1107  
Qy 684 DNGAAGKDRQIADIELLEINGOILYGRSHONASSIICAPSKVYIIFIRKNDANOMA 743  
Db 1108 APGGRADFHGCLKPGDRILSVNSVSLGVSHHALEITLONAEEDTLVLSQREKISKVP 1167  
Qy 744 VCPGNAVEPLPSNSENLOKETEPTVTSDAAVDLSSFPK-----VOHLELPK--DOGG 795  
Db 1168 STPVH---LTNEMKNYMKSS---YMODSAIDSSSKDHMSRGLRHIENSEFCPSGG 1219  
Qy 796 L--GLAISEEDLSGVIITKSLTEHGVAAIDGALKRGDQILAVDELIVGPIEKFISLKL 853  
Db 1220 LREGSLSSQDSRTESASLSQOVNGFFAS---HLGDQW---QESQHGSPSPSVIS-- 1269  
Qy 854 TAKMTVKLTIAHNPDSQAVPSAAGAAGEK-----NSSQSLAVPOSGPEPS 903  
Db 1270 --KATEKETFTSNOSKTKKPGISDVTYSDRGDSMDAATYSSSDHOTPKO---ESS 1324  
Qy 904 INTSRSTPAIFASDPATCPPIPGCETIIEISKRTGLSIVGSDTL--GAFIIE 961  
Db 1325 SVNTSMKNKXTEFSSPPK---PGDIFVELAKNDNSGIVTGCVNTVSRHGGIYVKA 1380  
Qy 962 VYEEGACCKDGLMAGDQILEVNGIDLKRAIIDEALINVRQTPORVRLTYIDEAPYKE 1021  
Db 1381 VIPOGAESDGRIRHKGDRVLAVNGVSLGATHKQAVETLRNGOVVHLLLEGQSPSTKE 1440  
Qy 1022 EV---C-----DTLIELOKKRGKGLIS-- 1042  
Db 1441 HVPVTPQCLTSQONAGGCEPEYKKTQYKDYSPYTEETFEKFLKN--SSGLGSFSRE 1499  
Qy 1043 ---IYKRNDCVFPVSDIYKGIADPDRILIOGDOILLVNGEDVRNASOEAVALKCSL 1099  
Db 1500 DNLIEQINASIVRVKKLKFAQGPAAESGKIDVDYLKYNKAGSLKGLSQGEVYSLRGTA 1559  
Qy 1100 GTVTELEVGRIRKG---PHESE----- 1117  
Db 1560 PEVFLLCRPPGVLPLEIDTALLPLQSPAQVLPNSSKDSQSPCEVSTSDENEMSDK 1619  
Qy 1118 ---RRPSOTQVSEGS-----LSPTF-----PL 1138  
Db 1620 SKKQCKSPKRSYSDSSGSDDLVTAPANISNTWSALHQTLSNMYSQAOSSHAEKP 1679  
Qy 1139 SGGST-----SSLESSK----- 1152  
Db 1680 SOEDTICTMFWYPPQKIPNKPREFEDSNPDLPPDMAFGQSYQPSSESASSMDKYHIHI 1739  
Qy 1153 ---KNALASITQGLR-----TYEMKKGPIDSLSIAG--VGSPLGVP 1192  
Db 1740 SEPTROENMTPKNDLENLEDEFEVELLITLISEKASLSLFTYKGNORIGCVYHVD- 1798  
Qy 1193 IFIAMHPTGVAOQOKLEVRGRIYITGSTECHTHQAVNLKNASISIEQVAVAGD 1252  
Db 1799 ---IOP---AKSDGRKPGDRILKVNDDVTNMTHDAVNLRAASTVRL----- 1844  
Qy 1253 VSVVTGHHQE-PASSLSFTGLTSTISIFODDGLPQCKSITLERGPDGIFSVIGYGSP 1311  
Db 1845 ---VIGRVLPRIPML-----PHLLPITLTCKNEELGFSICGHDLSL 1885  
Qy 1312 HDDLPIYKTVAKGASDGLKRGDQILIANVNGSLBEVTHEEA-----VAILKR 1362  
Db 1886 Y--QVYVSDIPRSVAALIEGNLQDLIDVHYHNGVSTQGMTEEVNRALDMSLPSLVLKA 1943  
Qy 1363 TKGVTLM 1370  
Db 1944 TRNDLPV 1951

RESULT 4  
US-08-596-291-3



ADDRESSEE: WOLF, GREENFIELD & SACKS, P.C.  
 STREET: 600 ATLANTIC AVENUE  
 CITY: BOSTON  
 STATE: MASSACHUSETTS  
 COUNTRY: USA  
 ZIP: 02210  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patentin Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/100,804  
 FILING DATE:  
 CLASSIFICATION:  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/596,291  
 FILING DATE: 09-AUG-1996  
 APPLICATION NUMBER: US 08/115,573  
 FILING DATE: 01-SEP-1993  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: PCT/US94/09943  
 FILING DATE: 01-SEP-1994  
 ATTORNEY/AGENT INFORMATION:  
 NAME: GATES, EDWARD R.  
 REGISTRATION NUMBER: 31,616  
 REFERENCE/DOCKET NUMBER: L0461/7003  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 617-720-3500  
 TELEFAX: 617-720-2441  
 TELEX:  
 INFORMATION FOR SEQ ID NO: 3:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 2465 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-09-100-804-3

Query Match 7.1%; Score 498; DB 3; Length 2465;  
 Best Local Similarity 21.9%; Pred. No. 2,4e-31;  
 Matches 252; Conservative 150; Mismatches 389; Indels 362; Gaps 41;

510 ELONNAYMNMNRRVELMREPSKSLGISTVGRMGSRSLNG----- 552  
 868 QLOMRAROSNODQIERASFRSLNLQAESVGNMGRALISTGLASSTLNTKLAVPLISV 927  
 553 --EVARGIFIKHV-----LEDSPAGKNGTL--KPGD--RIVEAPSOSESEPEKAPLCS 599  
 928 QAEILKRLSCSLSLYQPLQNSSKEKNKASWEKPREMSKSYHDSQSLYPHKKNVIV 987  
 600 VPPPPSAFAEM-GSDHTOSSAS-----KISQDVKED 631  
 988 NMEPPQVYAEIVKRPQMSRSDAESLAGVTKLNNSKSVASLNSRPERKHESDSSSIE 1047  
 632 EFG--YSKNKTRERYCITLG---ELHMIETELK-GHSGLGLS-LAGNKRDRSRSVITGID 684  
 1048 DPGQAYVLDVLRKRWISYSPERETTLVNLKRDARYGLGFOIIGGEKMETDGIETISVA 1107  
 685 PNGAGKDGKDRLOIADELLEINCOILYGRSHONASSIICAPSKVITIFIRNDANOMAV 744  
 1108 PGGPDPFGICGLKPGDRLLSVNSVSEGVSHMAIETLONAPDVLIVLSQPEKISKVPS 1167  
 745 CPGNAVEPLPSNSENLOKNETEPVTTSDAVADLSSFKN-----VOHLELPK--DOGGL 796  
 1168 TPVH-----LTNEMKNVYMKKS---YMODSAIDSSSKDHMGRTLRHISENSFGSGGL 1219  
 797 --GIAISEEDTLSTGYIISLTHGVAAATDGRKAKVDQILAVNDELIVGPIPIKFIISLKT 854  
 1220 REGSISSDQSRSTESASLSOSQVNGFPAS---HLGDQTV---QESOHGSPSPSVIS--- 1268  
 855 AKMTVKLTIHAEINPDQAVPSPAAGAAGEKK-----NSSOSLWVPOGSGSPEPEST 904

1269 -KATEKETFTDSNOSTKTKKPGISDVTDYSDRGDSMDENATYSSSDQHPKO---ESSSS 1324  
 905 RNTSRSSSTPAIPASDPANCPPIPGCEETIELSKRGTLGSLIVSGSDTL--GAFITHEV 962  
 1325 VNTSNKMNKFTSSSPK---PGDIFEYELAKNNSLGIYSTGVGNISVRNIGGIYNDV 1380  
 963 YEEGAACKDGRIMAGDQILEVNGIDLRKATHDEAINVLKOTPORVRLTYRDEAPYKEE 1022  
 1381 IPQGAESDGRTHKGDVYLVANGVSELEGATHKQAVETLRNTGOVYHLLLEKQSPSTKEH 1440  
 1023 V-----C-----DTLELQKKRPGKGLS----- 1042  
 1441 VFTPTQCTISDNOAGCGPEKKYKTTQYKDVSVTEENTFEVKLRN--SSGLGFSFRED 1499  
 1043 --YVGRNDTVGVFSDIYKGIADPDGRLIQDQILLVNGEDVRNASEAVALLKCSIG 1100  
 1500 NLIPQINASIVRYKKLFAGQPAESGKIDVDVILKVNAGSLKGLSQOEYISALRGAP 1559  
 1101 TYTLEVGRIKAG-----PRHSF----- 1117  
 1560 EVFLLICRPPEVLPRLPIDTALTPLQSPAQVLPNSKSDSOPSCVEQSTSDENEMSDKS 1619  
 1118 ---RRPSOTSGVSEGS-----LSSFTF-----PLS 1139  
 1620 KKQCKSPSRROSYSDSSGSGEDDLVTAPANISNTWSSALHOTLSNMVSOAOSHHEAPKS 1679  
 1140 GGSST-----SELSSESSK----- 1152  
 1680 QEDICTMFTYPOKIPNKEPEFEDSNPRLPRPMARGQSYQPQSEASASSMOKYHIHHS 1739  
 1153 -----KNALASELOGLR-----TVEKKRPTSLGISTINGG---VSGPLGDPVI 1193  
 1740 EPTROENMTPRLNDENLHDELEVEELLITLIKESKASLGFTYVKGNOICGYHVD-- 1797  
 1194 FIAMHPPTGVAQOTOKLRVGDRIYITICSTEGMHTGOAVNLKNASGSIEM-----QV 1247  
 1798 ---IDP---AKSDGRILKRGDLIKVNDIVYNTHTDAVNLBRAAKTYRLVIGYRRI 1851  
 1248 VAGDV-SVYTGHHQEPASSLSFTGLTSTSTIFDGLPPOCKSTTLERPGDLGFSYVG 1306  
 1852 TONTNVASFATGH-----KLNCNKEELGFSLCG 1879  
 1307 GYGSPHGLPIYVKTVPKGAASEDGRLKRGDQILAVNGOSLEGYTHEA-----V 1357  
 1880 GHDSLY--QVYIISDINPRSVATEGNLQLDLVIVYNGVSTGYMTLEEVNRALDMSLPS 1937  
 1358 AILKRTKGTITLM 1370  
 1938 LVLKATRNDLPVY 1950

RESULT 6  
 US-09-045-632-2  
 Sequence 2, Application US/09045632  
 Patent No. 6001575  
 GENERAL INFORMATION:  
 APPLICANT: HUGANIT, Richard L.  
 APPLICANT: DONG, Hualing  
 TITLE OF INVENTION: THERAPEUTIC USES OF GRIP AND  
 TITLE OF INVENTION: GRIP-RELATED MOLECULES  
 NUMBER OF SEQUENCES: 105  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
 STREET: 130 Water Street  
 CITY: Boston  
 STATE: MA  
 COUNTRY: USA  
 ZIP: 02109  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS

```

SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/045,632
FILING DATE: 19-MAR-1998
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/041,016
FILING DATE: 19-MAR-1997
ATTORNEY/AGENT INFORMATION:
NAME: Corleiss, Peter F.
REGISTRATION NUMBER: 33,860
REFERENCE/DOCKET NUMBER: 48147/1699-CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1112 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-045-632-2

```

```

Query Match      4.98; Score 340; DB 3; Length 1112;
Best Local Similarity 22.6%; Pred. No. 6,4e-19;
Matches 173; Conservative 90; Mismatches 265; Indels 238; Gaps 22;

```

```

OY 781 FKNVHLEPRDQ-GGIAISEEDLSGYI-----IKSLTEHGVAAATGRKVGDOI 832
DB 48 FKGSTVELMKREBTTLGCG-----TVSGGIDKDGKPRVSNLRGGIAASDOLDVGYI 101
OY 833 LAVDEIVGVPIEKFTSLTKAKMTVKLTHAENPDQAVPSAAGAAGEKKNSQSLM 892
DB 102 KAVGINALAKFRHDEILSLKNVGERVYLEVEYELP----- 137
OY 893 VPOSGSPESIRNTSSSTPAIFASDPATCPIIPGCEETIEISKRTG--LGLSIYVGS 950
DB 138 -----PVSIQSS-----VME-----RTVEVTLHKEGNTFGFVIRGGA 170
OY 951 D---TLGATLHIEVYEGAAKDGRLMAGDOIIEVNGIDLRKATHDEATVNLQTPQRY 1007
DB 171 HDRNRKSNPVVITCVPRGPDREGITKPGDRLLSVGIRLLGTHAEAMSLKQCGEA 230
OY 1008 RLTL-YRDEAPYKEEVEDLTIELOKKPKGGLSIYVK--RNDTGVSVDIYKGGIAD 1064
DB 231 TLLEIVDSANDSVATASGPLLEVAKTPGASLGVALTTSVCKKQYIVIDKISASIAID 290
OY 1065 PDGRLIOGDOIILVNGEDVRNASOEAAVALIKSLGIVTLEFV-----GRIK-AGPFH-- 1115
DB 291 RCGALHYGDHILSIDGTSMEYCTLAETQFLGNTTDQVKLEILPHQTRILAKGPDHVKI 350
OY 1116 -----SERRPSQ 1122
DB 351 QRSRQRLPMDPMASSQCSVHTNHHNHPHDHCVPALGFKALTPNSPPAMVSSSSSTTS 410
OY 1123 TSOVSEGLSSTFPPLSGSSST-----ESLESSKKNALASEIQTGRTV 1166
DB 411 MSAVSLSLNNGTLPRSLYSTSPRGTMWRRLKKKDFKSSLSLSTVGLAGVYHETT 470
OY 1167 E--MKKPTDSLGISINGV--GSPLGDPVPIFIAMMPTGVAAGTOKLAVGDRIYTTIGTS 1223
DB 471 EYVLTADPVYGFGLQIGSVFATETLSSPLISTYEDSPERCGVLIQIGRVAAINGIP 530
OY 1224 TEGTHTHOAVNLKSNAGSS-----IMQYVAGSDVSVVGHQEPASSSLSFTGLNST 1277
DB 531 TEDSTFEANOLDLRDSSITSKVLIEIEFDVAESVIPSSSGTFHVKLPKKHVEL-GITISS 589
OY 1278 IFDDDLGP-----QCR--- 1289
DB 590 PSSRKPDPLVTSIDIKKGSVAHRTGTLELDKLLAIDNIRLSDCSMEDAVQIILQCCEDLV 649

```

```

OY 1290 -----SITLER--GPDGLGFSIVGYSGSPHGLPIYVYTFYAK 1325
DB 650 KLIKIRKDEDNSEDQESSGAIYVELKRYGGPLGITIS-----GTEPFDPIIISLTKG 704
OY 1326 GAASEDRKLRGDDIITAVNGSLGCVTHEEAVALKRTKGTVTLMV 1371
DB 705 GLAERTGAIHIGDRILAIINSSSLKGPLSEIDHILLOMAGETVTLKI 750

```

# RESULT 7

```

US-09-045-632-3
Sequence 3, Application US/09045632
Patent No. 6001575
GENERAL INFORMATION:
APPLICANT: Huganir, Richard L.
TITLE OF INVENTION: THERAPEUTIC USES OF GRIP AND
TITLE OF INVENTION: GRIP-RELATED MOLECULES
NUMBER OF SEQUENCES: 105
CORRESPONDENCE ADDRESSES:
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
STREET: 130 Water Street
City: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/045,632
FILING DATE: 19-MAR-1998
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/041,016
FILING DATE: 19-MAR-1997
ATTORNEY/AGENT INFORMATION:
NAME: Corleiss, Peter F.
REGISTRATION NUMBER: 33,860
REFERENCE/DOCKET NUMBER: 48147/1699-CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1112 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-045-632-3

```

```

Query Match      4.98; Score 340; DB 3; Length 1112;
Best Local Similarity 22.6%; Pred. No. 6,4e-19;
Matches 173; Conservative 90; Mismatches 265; Indels 238; Gaps 22;

```

```

OY 781 FKNVHLEPRDQ-GGIAISEEDLSGYI-----IKSLTEHGVAAATGRKVGDOI 832
DB 48 FKGSTVELMKREBTTLGCG-----TVSGGIDKDGKPRVSNLRGGIAASDOLDVGYI 101
OY 833 LAVDEIVGVPIEKFTSLTKAKMTVKLTHAENPDQAVPSAAGAAGEKKNSQSLM 892
DB 102 KAVGINALAKFRHDEILSLKNVGERVYLEVEYELP----- 137
OY 893 VPOSGSPESIRNTSSSTPAIFASDPATCPIIPGCEETIEISKRTG--LGLSIYVGS 950
DB 138 -----PVSIQSS-----VME-----RTVEVTLHKEGNTFGFVIRGGA 170
OY 951 D---TLGATLHIEVYEGAAKDGRLMAGDOIIEVNGIDLRKATHDEATVNLQTPQRY 1007

```

```

Db 171 HDDRNRKRVVITCVYRPGGDDREGTKRGRLLSVDIRLLGTTTHAEMSLKQCGQEA 230
QY 1008 RLTL-YRDPAPKEEVCTLTLELOKKRKGLSIYVK--RNDTGVVSVIWKGLAD 1064
Db 231 TLLIEDVAMSQVATASCPRLVEVAKTPGASLGVALLTTSVCCNOVIYIDIKSASIA 290
QY 1065 PGRLIOGQDILLVNGEDVRNASQEAVALLCISLGTVLEV-----GRIK-AGPFI--- 1115
Db 291 RCGALHVGDIHLSIDGTSMEYCTLAETATPFLGNTDQVLELPHHQTLALKGDPHVXI 350
QY 1116 -----SERRPSQ 1122
Db 351 QSRDRLPMDPWASSQSVHTNHNHPHDPCHVPALGFPKALTPNSPRAWSSSSPTS 410
QY 1123 TQVSESGISSTFPPLSGSSTS-----ESLSSSKKNALASEIOGLRTV 1166
Db 411 MSAVSSLSLNMGTLPRLSTLSTSPRGTMRRRLKKDKFSSLSIASVTGLAQVHTET 470
QY 1167 E--MKKGPDSIGSIAGV--GSPLDVPIFTAMMHPGTVAOTOKLRVGRIRVTICGS 1223
Db 471 EYVLRADPTGFGIOLOGSVFATEPLISPLISYIFADSPAERCGLVQIGDVMALNGIP 530
QY 1224 TEGMTHQAVNLKNASGS-----TEMQVAGDVSVYTGHNQBPASSLSFTGLTSTS 1277
Db 531 TEDSTFEENQOLLRDSSITSKVTLIEFDVAVSVIPSSGTFHVKLPRKHSVEL-GITISS 589
QY 1278 ITODDLGP-----OCK--- 1289
Db 590 PSSRRPGDLVISDIKKSVAHRTGTLELDKLLAIDNIRLDSCEMEDAVOIQOCEDLV 649
QY 1290 -----SITLER--GPDGLGFSIVGSGPHGLDPIYVTKVFAK 1325
Db 650 KLIKIKEDNSDEQESSGAIITYVELKRYGGLGITIS-----GTPEPDPPIISSLTNG 704
QY 1326 GAASEDRKLKRGDQIIAVNGQSLGVTHEEVAVALIKRTKGVTLAV 1371
Db 705 GLAERTGAIHGDRILAINSSSLKGRPLSEDIHLLQMAGETVTLKI 750

```

## RESULT 8

```

US-09-045-632-15
; Sequence 15, Application US/09045632
; Patent No. 6001575
; GENERAL INFORMATION:
; APPLICANT: Huganir, Richard L.
; TITLE OF INVENTION: THERAPEUTIC USES OF GRIP AND
; TITLE OF INVENTION: THERAPEUTIC USES OF GRIP AND
; NUMBER OF SEQUENCES: 105
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/045,632
; FILING DATE: 19-MAR-1998
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/041,016
; FILING DATE: 19-MAR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Corless, Peter F.
; REGISTRATION NUMBER: 33,860
; REFERENCE/DOCKET NUMBER: 4817/1699-CIP
; TELECOMMUNICATION INFORMATION:

```

```

; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 702 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-045-632-15

```

```

Query Match 4.8%; Score 335; DB 3; Length 702;
Best Local Similarity 22.5%; Pred No. 7.3e-19;
Matches 171; Conservative 90; Mismatches 261; Indels 238; Gaps 22;

```

```

QY 787 LELPKDQG-GLGIAISEEDTLGVI-----IKSLTEHVATDRLKRVGQDIIAVNDE 838
Db 3 VELMKKEGTTLC-----TVSGGIDKDGKPRVSNLRGGIARSDQDLVDGYIKAVNGI 56
QY 839 IYVGPPIEFISLTKAKTVKLTTHAENPDQAVPSAAGAASGEKNSQSLAMPQSGS 898
Db 57 NLAKEFHDEIISLKNVGRVYLEVEYELP----- 86
QY 899 PEPESIRNTSRSTPAIFASDPATCPPIPGCEFTIEISKGRNG--LGSLVGGSD--TL 953
Db 87 --PVSLQSS-----VMF-----RTVEVTLHKGNTPFGVYIRGAHDDRK 125
QY 954 LGAFITHEVEYEGACAKDGRMLWAGQILEVNGIDLRKATNDEAIVNLTQTPQVRVLTLY 1012
Db 126 SRPVVITCVYRPGGDDREGTKRGRLLSVDIRLLGTTTHAEMSLKQCGQEA 185
QY 1013 RDEAPRKEEVCTLTLELOKKRKGLSIYVK--RNDTGVVSVIWKGLAD 1070
Db 186 DVSAMDVATASCPRLVEVAKTPGASLGVALLTTSVCCNOVIYIDIKSASIA 245
QY 1071 QCDQDILLVNGEDVRNASQEAVALLCISLGTVLEV-----GRIK-AGPFI--- 1115
Db 246 VGDHILSIDGTSMEYCTLAETATPFLGNTDQVLELPHHQTLALKGDPHVXI 305
QY 1116 -----SERRPSQ 1122
Db 306 LPMDPWASSQSVHTNHNHPHDPCHVPALGFPKALTPNSPRAWSSSSPTS 365
QY 1129 GSLSFTFPPLSGSSTS-----ESLSSSKKNALASEIOGLRTV--MKK 1170
Db 366 SSLNMGTLPRSLSTSPRGTMRRRLKKDKFSSLSIASVTGLAQVHTET 425
QY 1171 GPTDSLIGISAGV--GSPLDVPIFTAMMHPGTVAOTOKLRVGRIRVTICGTEG 1229
Db 426 DPTVGTGLOGSVFATEPLISPLISYIFADSPAERCGLVQIGDVMALNGIP 485
QY 1230 TQAVNLKNASGS-----TEMQVAGDVSVYTGHNQBPASSLSFTGLTSTSIF 1283
Db 486 EKANQLLRDSSITSKVTLIEFDVAVSVIPSSGTFHVKLPRKHSVEL-GITISS 544
QY 1284 GPP-----OCK--- 1289
Db 545 GPDVLVISDIKKSVAHRTGTLELDKLLAIDNIRLDSCEMEDAVOIQOCEDLV 604
QY 1290 -----SITLER--GPDGLGFSIVGSGPHGLDPIYVTKVFAK 1331
Db 605 DEEDNSDEQESSGAIITYVELKRYGGLGITIS-----GTPEPDPPIISSLTNG 659
QY 1332 GRLKRGDQIIAVNGQSLGVTHEEVAVALIKRTKGVTLAV 1371
Db 660 GAIHGDRILAINSSSLKGRPLSEDIHLLQMAGETVTLKI 699

```

## RESULT 9

```

US-09-045-632-16
; Sequence 16, Application US/09045632
; Patent No. 6001575

```

GENERAL INFORMATION:  
APPLICANT: Huganir, Richard L.  
TITLE OF INVENTION: THERAPEUTIC USES OF GRIP AND  
NUMBER OF SEQUENCES: 105  
CORRESPONDENCE ADDRESS:  
DIREK, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/045,632  
FILING DATE: 19-MAR-1998  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/041,016  
FILING DATE: 19-MAR-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Corless, Peter F.  
REGISTRATION NUMBER: 33,860  
REFERENCE/DOCKET NUMBER: 48147/1699-CIP  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 16:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1018 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-045-632-16

Query Match 4.88; Score 335; DB 3; Length 1018;  
Best Local Similarity 22.55; Pred. No. 1.4e-18;  
Matches 171; Conservative 90; Mismatches 261; Indels 238; Gaps 22;

787 LELPRDQ-GGIAISEEDTLGVY-----IKSLPEHVAATDGRKVDOLLAVDE 838  
3 VELAKKEGTTIGC-----TVSGIDKDKGRVSNLROGGIAANSDDLDVGDYIKAVNGI 56  
839 IVGVPIEKFISLTKAKMTVKLTIHENPDQAHPVPSAAGASEKKNSSQSLVWPOSQS 898  
57 NLAKFRHDEIISLTKNKGVERVLEVEYELP----- 86  
899 PEPSIRNTSSSPALFASDPATCPPIIPGCEITTEISKGTG--LGLSTVGGSD--TL 953  
87 --PVSIGSS-----VWF-----RTVEVTLHKEGTYTFGVINGAHDDRKNK 125  
954 LGATIHVEYEGAAKGRGLMAGDQILEVNGIDLRKATHEAIVNROTPORVRLTY 1012  
126 SRPVYITCVRRGGPDGRGTTIKPDRLSLVDGIRLLGTTHEANSLIKQCCQERTLLLEY 185  
1013 RDEAPYKEEVEDPRLTELQKPKGGLSLVGR--RNDTGVFVSDIVKGIADPDGRLI 1070  
186 DVSAMDVATASGPLVAVAKTPGASLGVALTVSCCKKQIVYIDKIKSASIAIRGALH 245  
1071 QGDDILLVNGEDVARNASOEAVALLKCSLGVTLV-----GRIT-AGPRH----- 1115  
246 VGDHILSIDGTSMEYCTLAETQFLAGTTDQVKLEILPHHQTRIALKGPDHVKIORSDRQ 305  
1116 -----SERPSQTSQVSE 1128  
306 LPMPPMASSQCCVHTNHHNHPHDPHCHVPALGPALKALTNSPPAMVSSSSPTSMASVSL 365

1129 GSLSFTEPLSGSSTS-----ESLESSKKNALASEIGLRTVE--MKK 1170  
366 SLSMNGTIPRLSLYSTSPRGTMRRRLKKDKSSLSIASSTVGLAGOVHETTEVILTA 425  
1171 GPTDSLGISIAGV-GSPFLGDPVPIFIAMHPTVAAOTKLRVDRIVTTCGSTEGMTA 1229  
426 DPVTCFGILOGSVFATETLSSPLISYIEADSPAEKRGVLIQIDRVAMINGIPTEDSTF 485  
1230 TVAVNLKNAGS-----IEMOVYAGDVSVYGHNOEPASSLSLSTGLSTSTIFDDDL 1283  
486 EEAQQLNRDSSITSKYTLEIFDVAESVIPSSGFPHVLPKRHSVEL-GTITSSPSSRRK 544  
1284 GPP-----OCK----- 1289  
545 GDPVLVSDIKKGSVAHRTGLELGDKLAIIDNRDSCSMEDAVOIIQCCEDLVKLRK 604  
1290 -----STTLER--GPDGLGFSIVGVGSPHGDLPYVKTVFANGAASED 1331  
605 DEDNSDEQESSGALITVVELKRYGPGIGTIS-----GTEEPFDPIITISLTKGLAERT 659  
1332 GRLKRGDIIIVNGOSLEGVTHEEVAVALKRTKGTVTLMV 1371  
660 GAHIGDRILAINSSSLKGPLSEDIHLLOWAGETVTLKI 699

RESULT 10  
US-09-045-632-32  
Sequence 32, Application US/09045632  
Patent No. 6001575  
GENERAL INFORMATION:  
APPLICANT: Huganir, Richard L.  
TITLE OF INVENTION: THERAPEUTIC USES OF GRIP AND  
NUMBER OF SEQUENCES: 105  
CORRESPONDENCE ADDRESS:  
DIREK, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/045,632  
FILING DATE: 19-MAR-1998  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/041,016  
FILING DATE: 19-MAR-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Corless, Peter F.  
REGISTRATION NUMBER: 33,860  
REFERENCE/DOCKET NUMBER: 48147/1699-CIP  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 32:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1061 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-045-632-32

Query Match 4.88; Score 335; DB 3; Length 1061;

Best Local Similarity 22.5%; Pred. No. 1.5e-18;  
Matches 171; Conservative 90; Mismatches 261; Indels 238; Gaps 22;

```

OY 787 LELPKDQG-GIGAISEEDTSLGVI-----IKSLTEHVAATDGRKLVGDQILAVDE 838
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 3 VELMKKEGTTLC-----TVSGGIDKDGKPRVSNLRQGIARSDQDLVDYIKAVNGI 56
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 839 IYVGPRIEFILTKAKTVLTIHAENPDQAVPSAAGAAGEKNSQSILMPQSS 898
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 57 NIAKFRHDEIISLKNVGRVVELEYELP----- 86
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 899 PEPEIRNTRSRSTPAIFASDPATCPRIIGCETTEISKRTG--LGLSIVGSD---TL 953
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 87 --PVSLQSS-----VMF-----RTVEVTLHKEGNTFGFVIRGAHNRK 125
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 954 ICAFIIIEYEEGACKDGRLNAGQILEVNGIDLRKATHEAIVNLQTPORVRLTL-Y 1012
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 126 SRPVYITCVRPGPDRECTIKPRRLISVDGIRLIGTTHAEMSLKOCGGEATLLIEY 185
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1013 RDEAPYKEEVCDTLTIELQKKRPGKGLSIVGK--RNDTGVFVSDIVKGIADPDGRLI 1070
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 186 DVSANDSVATAGPLVEAKTPGASLGVALLTVCCNKQVYIDKIKASIAIDRCGLH 245
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1071 QGDQILLVNGEDVRNASEPAVALLKSLGTVLEY-----GRK-AGPER----- 1115
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 246 VGDHILSIDGTSMEXCTLAETQFLGNTTDQYKLEILPHHQRRLAKGPDHVKIQRSNQ 305
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1116 -----SRRPSQTSQVSE 1128
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 306 LPMDPWASSQCSVHTNHNHPPHDPCHRVAPALGPRKALTRNPPRANVSSSPMSAYSL 365
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1129 GLSLSTFPLSGSSTS-----ESLESSKKNALASEIGLRTVE--MKK 1170
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 366 SSLNMGTLPRSLXSTSPRGTMMRRRLKKKDFSSLSLASTYGLAQVYHTEETEVYLA 425
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1171 GPTDLSIGSIAGV--GSPGLDVPFIAMMPTGVAQOTKLRVGRIVTICSTEGMTN 1229
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 426 DPTVGTGIDQSGVFATETLSSPRLSYLEADSPARCVQLQIGRVMAINGIPEDSIF 485
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1230 TQAVNLKNAAGS-----IEMQYVAGGVSVYTGHHQBPASSLSFTGLTSTISITQDL 1283
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 486 EKANOLRDSITSKYVTLIEFDVASEVIPSQGFHVKLPRKHSYVL-GITISSPSRRP 544
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1284 GPP-----OCK----- 1289
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 545 GPPVYISDIKKSVAHRTGTLELGRKLLIDNIRLDSCEMEDAVOILQCCEDLVKLKIK 604
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1290 -----SITLER--GPDGLGFSTVGGYSGPHGDLPIYVKTVEFAKGAASED 1331
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 605 DEDNSEQSSGALITVELKRGPGIGITIS-----GTEEPDPITLISLTKGLAERT 659
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1332 GRLKRGDQIIAVNGOSLEGVTHEEVAAILKTKRGVTLMV 1371
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 660 GAHIGDRILAINSSSLKGPLSEDIHLQMGFTVTLKI 699
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 11  
US-08-410-804-1  
Sequence 1, Application US/08410804  
Patent No. 5632994

GENERAL INFORMATION:

APPLICANT: Reed, John C.  
APPLICANT: Sato, Takaaki  
TITLE OF INVENTION: FAS ASSOCIATED PROTEINS  
NUMBER OF SEQUENCES: 22  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Cathryn Campbell  
STREET: 4370 La Jolla Village Drive, Ste 700  
CITY: San Diego  
STATE: California  
COUNTRY: United States  
ZIP: 92122  
COMPUTER READABLE FORM:

```

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/410.804
FILING DATE: 27-MAR-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/259,514
FILING DATE: 14-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn
REGISTRATION NUMBER: 31,815
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 610 amino acids
TYPE: amino acid
MOLECULE TYPE: protein
US-08-410-804-1

Query Match 4.8%; Score 334.5; DB 1; Length 610;
Best Local Similarity 23.0%; Pred. No. 6.3e-19;
Matches 155; Conservative 72; Mismatches 221; Indels 225; Gaps 21;

OY 842 GYPIKEFISLTKAKTVLTIHAENPDQAVPSAAGAAGEK-----NSSQSL 891
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2 GSPDSVVIS-----KATEETITDSNQSTKPRGISDVTYDSDRGSDMDATYSSQDH 56
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 892 WYPOSQPEPEIRNTRSRSTPAIFASDPATCPRIIGCETTEISKRTGLSIV---- 947
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 57 QPQPKQ---ESSSVNTSNMKNFTSSPPK---PGDIFVELAKNDNSLGSIVTVLED 109
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 948 -GGSDDL--GAFIIIEYEEGACKDGRLNAGQDILEVNGIDLRKATHEAIVNLQTP 1004
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 110 KGVNTSVRHGGIYVAVYVPOGAESDGRIRHKGRVLAENVGSLGATHKQAVETLRNG 169
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1005 QVRLTLVDEAPYKEEV-----C-----DLETLI 1029
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 170 QVHLLERQSGSTSKEHVPVTPQCTLSQDQNOAGCPKPVKTTQVKDYSEVTEETFEV 229
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1030 ELQKPKGKGLS-----IYKRNQDGVFVSDIVKGIADPDGRLIOGQILLVNGED 1082
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 230 KLEFN--SSGLGSFSEKEDLIRBQINASTVRYKKLFPQGPAAESKIDVGDVILKNGAS 288
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1083 VRNASQEAVALKKSGLGTVTLEVGRIKAGP-----HSRRPSQT 1123
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 289 LKGLSQOEIVASALRGTAPEVFLLCRPPGVLPETIDTALLPLQSPAVLPNNSKDSQP 348
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1124 SQVSEGLSSTFTPLSGSSTSESLESSKKNALASIQGLRTVEKMGKPTDLSIGSIAG 1183
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 349 SCVEQ-----STSSDENEMDSKQOC-----KSPSRDSYSDSSG 384
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1184 VSPPLGDVPFIAMMPTGVAQOTKLRVGRIVTICSTEGMTHTQAVNLKNGASGI 1243
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 385 SGE-----DGLVT-----APAINSNSTWS 404
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1244 EMQYVAGGVSVYTGHHQBPASSLSFTGL-----TSTSIPODDLCP--- 1285
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 405 ALHQTLNNVSOAOSHHEAPKQOEDTICIMFYYPQKIPRKPEEDSNPSPPLPDMAFGS 464
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1286 --PQCKS-----ITLSPGPG 1299
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 465 YOPQSESSASSMDKYYHHIISEPTROENMTPLKNDLENHLDPELEVLLTTLKSENG 524
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1300 -LGFSIVGGYSGPHGDLPIYVKTVEFAKGAASEDGRKRGDQIIAVNGOSLEGVTHEEVA 1358
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

Db 525 SLGFTVTKG---NORIGCYHVDYI-QDPAKSDRLKPGDRILKYNDDTNTMHTDAVN 579

QY 1359 ILKRTKGTVMV 1371

Db 580 LLRAASKTVRLVI 592

RESULT 12

US-08-259-514-1

Sequence 1, Application US/08259514

Patent No. 5747245

GENERAL INFORMATION:

APPLICANT: Reed, John C.

APPLICANT: Sato, Takaaki

TITLE OF INVENTION: FAS ASSOCIATED PROTEINS

NUMBER OF SEQUENCES: 22

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Cathryn Campbell

STREET: 4370 La Jolla Village Drive, Ste 700

CITY: San Diego

STATE: California

COUNTRY: United States

Zip: 92122

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/259,514

FILING DATE: 14-JUN-1994

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Campbell, Cathryn

REGISTRATION NUMBER: 31,815

REFERENCE/DOCKET NUMBER: P-LJ 9954

TELECOMMUNICATION INFORMATION:

TELEPHONE: (619) 535-9001

TELEFAX: (619) 535-8949

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 610 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-259-514-1

Query Match 4.88; Score 334.5; DB 1; Length 610;

Best Local Similarity 23.08; Pred. No. 6.3e-19;

Matches 155; Conservative 72; Mismatches 221; Indels 225; Gaps 21;

QY 842 GYPIKEFTSLKTKAMTVKLTIIHAENPDQAVPSAAGASGEK-----NSSQSL 891

Db 2 GSPSPSVIS-----KATKEFTSDNSOSKTKKPGISDVTDSGDSDMDATYSSSDH 56

QY 892 MWPGSGSPESINNTSSSTPAIFASDPATCPIIPGCEITIEISKGTGIGLSIV---- 947

Db 57 QTPKQ-----ESSSVNTSKNMFKTFSSSPK-----PGDIFEVLAKNDNSLGISVTLFD 109

QY 948 -GGSDDL--CAFTIHEYEGGACAKGRLMAGQILEVNGIDLRKATHDAIVLQTP 1004

Db 110 KGVNTSVRHGIVYKAVIPQGAESDGRIRHGDRLVAVNVSLEGATHKQAVETLRNTG 169

QY 1005 QRVRLTIYRDAPYKEEV-----C-----DPLTI 1029

Db 170 QVALLLEKGGSPISKEHVPTPCCTISDONAGCGPEKAKKTVQVVDYVTEENFEV 229

QY 1030 ELAKPKGGLGLS-----IVGRNDTGVFSDIVKGIADPDGRLIQCQDILLVNGED 1082

Db 230 KLFKN-SSGIGFSFSRBDNLPEQINASIYKVKLLFGPGPAESGKIDVGVILKVGAS 288

QY 1083 VRNASEAVALLKCSIGTYTLEVGRIKAGP-----HSERRPSQT 1123

Db 289 LKGLSQGEVIALKRGTAPEVFLLCRPPGVLPEDIDRALLPLOSPPAQVLPNSSKDSQP 348

QY 1124 SQVSEGLSFTFPLSGSSSTSESSSKKNALASETIOGLFTVEMKGPIDSIGSIAGC 1183

Db 349 SCVEQ-----STSDENEMSDSKKQC-----KSPSRDSYSSSG 384

QY 1184 VGSPLGDVPIFIAMHPTGVAAGQTKLRVGDRIYTGSTEGMTHQAVNLKNAGSI 1243

Db 385 SGE-----DGLVT-----APANISNTWS 404

QY 1244 EMQVYAGDVSVYVTHQEPASSLSFTGL-----TSTSIPODLGP--- 1285

Db 405 ALHQTLNMTVSAQASHHAPKQSDICTMFTYPOKIPNKPEDSNPSPLPDMARGOS 464

QY 1286 --POCKS-----ITLERPGD 1299

Db 465 YQPOSESASSMDKIHIIHISEPTROBNWTPKNDLENHLEDELEVELLITLIKSG 524

QY 1300 -LGFSTVGYGSPHGDLPYIKTVFAKGAASEDRKRGDIIAVNGQSLGVTHEEAVA 1358

Db 525 SLGFTVTKG---NORIGCYHVDYI-QDPAKSDRLKPGDRILKYNDDTNTMHTDAVN 579

QY 1359 ILKRTKGTVMV 1371

Db 580 LLRAASKTVRLVI 592

RESULT 13

US-08-858-311-1

Sequence 1, Application US/08858311

Patent No. 5876939

GENERAL INFORMATION:

APPLICANT: Reed, John C.

APPLICANT: Sato, Takaaki

TITLE OF INVENTION: FAS ASSOCIATED PROTEINS

NUMBER OF SEQUENCES: 22

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Cathryn Campbell

STREET: 4370 La Jolla Village Drive, Ste 700

CITY: San Diego

STATE: California

COUNTRY: United States

Zip: 92122

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/858,311

FILING DATE:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Campbell, Cathryn

REGISTRATION NUMBER: 31,815

REFERENCE/DOCKET NUMBER: P-LJ 1389

TELECOMMUNICATION INFORMATION:

TELEPHONE: (619) 535-9001

TELEFAX: (619) 535-8949

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 610 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-858-311-1





Db 593 FHVKLPKRRGVELGITTSSASRRKGEPLIISDIKKSVAHRTGTLPEQDKLLAIDNIRLD 652  
 QY 1085 NASOE-AVAALLKCSLGTVLEVGRIKAGPHERPSQTS--OVSEGLSSTFFPLSGS 1141  
 Db 653 HCPMEYAVQILPOCE-DLVKTKI-----RKEDNSDEDESSGAVS-----691  
 QY 1142 STSELESSSKKNMLASEIQLRTVEKK--GPTDSIGISIAGVGSPLGDPVPIFIAMH 1199  
 Db 692 -----YVELKRYGCP--LGITIS-GTEEPFD--PIIISGLT 723  
 QY 1200 PTGYAOTOKRLVGDRIYITIGTSTEGMTHQAVNLKNAGSIEMOV-----A 1249  
 Db 724 KRGLAERTGASMLGTAWPSXSVSLKGRPLSEALHLLQVAGEYTLKIKQDLRPLRQ 783  
 QY 1250 GGDVSVYTGHHQEPASSLSFTGLTSTISFODDLGPPQCKSITLERGPDGLGFSIV-GGY 1308  
 Db 784 SGSLSEASDVDEDEPPALKG--GLITTHF-----SPAVPSV--DSAVESGSSATEGCF 833  
 QY 1309 GSPHGDLPYKTYFAKGAASEDRKLR 1336  
 Db 834 GSGS-----YTPQVAVRSVTPQEWRRSR 857

## RESULT 15

US-09-045-632-49  
 : Sequence 49, Application US/09045632  
 : Patent No. 6001575  
 : GENERAL INFORMATION:  
 : APPLICANT: Huganir, Richard L.  
 : APPLICANT: Doong, Hualing  
 : TITLE OF INVENTION: THERAPEUTIC USES OF GRIP AND  
 : NUMBER OF INVENTION: GRIP-RELATED MOLECULES  
 : CORRESPONDENCE ADDRESSES:  
 : ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
 : STREET: 130 Water Street  
 : CITY: Boston  
 : STATE: MA  
 : COUNTRY: USA  
 : ZIP: 02109  
 : COMPUTER READABLE FORM:  
 : MEDIUM TYPE: Floppy disk  
 : COMPUTER: IBM PC compatible  
 : OPERATING SYSTEM: PC-DOS/MS-DOS  
 : SOFTWARE: Patent in Release #1.0, Version #1.30  
 : CURRENT APPLICATION DATA:  
 : APPLICATION NUMBER: US/09/045,632  
 : FILING DATE: 19-MAR-1998  
 : CLASSIFICATION: 435  
 : PRIOR APPLICATION DATA:  
 : APPLICATION NUMBER: US 60/041,116  
 : FILING DATE: 19-MAR-1997  
 : ATTORNEY/AGENT INFORMATION:  
 : NAME: Corleiss, Peter F.  
 : REGISTRATION NUMBER: 33,860  
 : REFERENCE/DOCKET NUMBER: 48147/1699-CIP  
 : TELECOMMUNICATION INFORMATION:  
 : TELEPHONE: 617-523-3400  
 : TELEFAX: 617-523-6440  
 : INFORMATION FOR SEQ ID NO: 49:  
 : SEQUENCE CHARACTERISTICS:  
 : LENGTH: 1050 amino acids  
 : TYPE: amino acid  
 : TOPOLOGY: linear  
 : MOLECULE TYPE: protein  
 : US-09-045-632-49

## Query Match

4.7%; Score 331.5; DB 3; Length 1050;

Best Local Similarity 21.8%; Pred. No. 2.9e-18;

Matches 202; Conservative 130; Mismatches 321; Indels 275; Gaps 42;

QY 525 VELKREPSKLSIGISIVGR-----GMSGR---LSNGEYWR---GIFIKHVL 564

Db 89 VELIKREGSTIGLTISGTDKDGKPRVSNLRPGGLAARSDLLANGDYIRSVNGIRLRLR 148  
 QY 565 EDSFAGKNGTL--KPGRIYAPASQSESEPEKALCSVPPPPAFAMGSDHQSASK 622  
 Db 149 HDEIT--TLKNGVERVY-----LEVYE-----LPPAP-----ENNRI 182  
 QY 623 ISODVDEDEFGYSWKNIIRERYGTLTGELHMIIELEKSHGSLAG----NKDRSMV 678  
 Db 183 ISKTVD-----VSLKENSNGFPLRGAMHDLKSRPLY 217  
 QY 679 FIVGIDPNAAGDKRLQIADLELINOIIGRSHQIA--SIITCAPSKYKIIFIRNKD 737  
 Db 218 -LTYVRPGPANREGSLKVGDRLLSIXGIPLHGASHATATLQOC-----SHE 265  
 QY 738 AVNMAVCPGNAVPELPDSNENLONKETEPTVTSDAVDLSFRKNOHELPRDQGLG 797  
 Db 266 ALFOV-----EYVATPDVYANASGLVYEIAKTPGSALGIS 302  
 QY 798 IAISEEDTLGCVIILKSLTEHGAATDGRKAVGQDILAVDEIVVGYPIEKELSLKTAKM 857  
 Db 303 LTTGSHRRKPAITIDRIKPAVVDKNGALHAGEHILAIIDGTSTEHCSLVEATKLLASYTE 362  
 QY 858 TVKLTIAHENDSQ---AVPSAGASGEKKNSSOLMVPQSGSPPESTR-----905  
 Db 363 KVRLEI-LPAQSRRLPKPEAVRIQRSEQLHNDP-CVPSCHSPRSHCAPTWAPGQ 420  
 QY 906 --NTSRSTPAIFASDPATCPIIPGCEITIEISG---RT-----940  
 Db 421 DOSRSVSTP---FSSPFPNAPF-CANASTLPFGPSPRTRARRORREHRSLSLA 476  
 QY 941 -----GLGSIYGG--SDTLGAFIIEVEEGAAKCD 971  
 Db 477 SSTVPGQGIYHTEFTEVVLGCDPLSGQLQLOGIFATELTLSPPLYVRFEPSPAERC 536  
 QY 972 GRWAGQIILEVNGIDLRKATHDEAINVLRQT--PQVRRLTYNDEAPYKEEV---CDT 1026  
 Db 537 GLLOVGDRLVAINGIATEDGTEQANOLLRDALARKIVLEIEFDVA---ESVYPSGT 592  
 QY 1027 LTIELKKPKGKGLSTI--VGRNDYGVFVSDIYKGIADPDGRLLIOGDOILLVNGEDVR 1084  
 Db 593 FHVKLPKRRGVELGITTSSASRRKGEPLIISDIKKSVAHRTGTLPEQDKLLAIDNIRLD 652  
 QY 1085 NASOE-AVAALLKCSLGTVLEVGRIKAGPHERPSQTS--OVSEGLSSTFFPLSGS 1141  
 Db 653 HCPMEYAVQILPOCE-DLVKTKI-----RKEDNSDEDESSGAVS-----691  
 QY 1142 STSELESSSKKNMLASEIQLRTVEKK--GPTDSIGISIAGVGSPLGDPVPIFIAMH 1199  
 Db 692 -----YVELKRYGCP--LGITIS-GTEEPFD--PIIISGLT 723  
 QY 1200 PTGYAOTOKRLVGDRIYITIGTSTEGMTHQAVNLKNAGSIEMOV-----A 1249  
 Db 724 KRGLAERTGASMLGTAWPSXSVSLKGRPLSEALHLLQVAGEYTLKIKQDLRPLRQ 783  
 QY 1250 GGDVSVYTGHHQEPASSLSFTGLTSTISFODDLGPPQCKSITLERGPDGLGFSIV-GGY 1308  
 Db 784 SGSLSEASDVDEDEPPALKG--GLITTHF-----SPAVPSV--DSAVESGSSATEGCF 833  
 QY 1309 GSPHGDLPYKTYFAKGAASEDRKLR 1336  
 Db 834 GSGS-----YTPQVAVRSVTPQEWRRSR 857

Search completed: July 12, 2001, 14:40:58  
 Job time: 201 sec

Fri Jul 13 15:00:07 2001

us-09-502-698-1.rai

Page 13





Interactions and is used for screening for proteins for use in treatment of cell proliferation disorders such as cancer

PS Claim 1; Page 84-87; 240pp; Japanese.

CC This sequence represents the mature portion of a new protein containing a PDZ domain encoded by the clone 38-2-1, whose expression in human umbilical vascular endothelial cells (HUVEC) is enhanced by stimulation with tumor necrosis factor (TNF) alpha. The new protein is used to identify proteins which bind to it (particularly to the PDZ domains) and the genes encoding them, for use in the treatment of cell proliferation disorders such as cancer.

CC disorders such as cancer.

Sequence 1005 AA:

Query Match 100.0%; Score 5085; DB 20; Length 1005;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1005; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MLQNVSKSEFRTINIAKNSLSGMTVSANKDGLMIVRSIIHGATSRDRIAGPCIL 60  
1 mlgvnskesfertlniakgnsslgmtvsankdglmivrsiihngatstrdriagpcil 60  
61 SINESTISVTNAQARAMLRRHSLIGPDIKITYVPAHELEEFKISLGQSGRNALDIFS 120  
61 sinestisvtnaqaramlrrhsligpdikityvpaheleefkislsgsgrnaldifs 120  
121 SYTRDIPELPEREGESELONTAVSMNORRVELMREPKSICISIVGRGMSR 180  
121 sytrdipelperegeeselnatavsmnorrvellmrepskislsgivgrgmshr 180  
181 LSNGEVMKGIKIKHVELESPAGKNGTLKPCDRIEAPSSQSESEPEKAPLCSVPPPSAF 240  
181 lsngemvngikihvleespagkngtlkpgdrievapsqsepekaplcsvpppsaf 240  
241 AEMGSDHTOSSAKTISQDVKEDDFGYSWKIRERYGTGLGELMIELEKSHSLGSLA 300  
241 aemgsdhtossaktisqdvkeddfgywknirerygtglgelmielekshslgsl 300  
301 GNRKSRMSVITVIGIDPGAAGKDRLOIADELLEINGQILYGRSHONASSITICAPSKV 360  
301 gnrksrmsvityigidpgaagkdrloiadelleingqilygrshonassiticapskv 360  
361 KIIFIRNKDAVNQAVCPGNAVEPLPSNSENLQNKREPTVTSDAAVDLSSFKNVOLLE 420  
361 kifirnkdaavnqavcpгнаveplpsnsenlqnkreptvttsdaavdlssfkvnv 420  
421 LPKDGGGLGIAISEEDTISGVTIKSLTEHGVAAFDGRKVGDDQLAVDEIVGPIEKE 480  
421 lpkdggglgiaiseedtsgvtyiksltehgvaafdgrrkvgddqlavdeivgpiek 480  
481 ISLTKTKMFWKLTITHAENPDSQAVPSAAGASEKKNSGSLWVPOSGSEPSINTS 540  
481 isltkktkmfwkltithaenpdsqavpsaagasekknsgslwvpsgspepsints 540  
541 RSTPAIFASDPATCPIIPGCERTIEISKGTGSLISVGGSDTLGAFIIEVYEEGAA 600  
541 rstpaifasdpatcpiipgcertieiskgtgslisvgsdltlgafiihevyeegaa 600  
601 CKDRLWAGDQIILEVNGIDLRKATHDEALINVLROTPORVRLTLKRDAPRYKEEVCOTLT 660  
601 ckdrlwagdqiilevngidlrrkathdealinvlrtpprvrltlkyrdaprykeevcot 660  
661 IELKKPKGKGLSIVGKRNDDGVFVSDIYVGAADPDGRLIOGDOLLVNGEDVRNASQ 720  
661 ielkkpkkgkglsvgkrrndgvfvsdiyvgaadpdgrrliogdollvngedvrnasq 720  
721 EAVAAALKCSLGTVTEVGRKAPFHSERRPSQTSQVSEGLSSFFPPLSGSSTSESL 780  
721 eavaallkcslgtvtevgrrkapfhserrpsqtsqvseglssffpplsgsstse 780  
781 SSKKNMALASEITQGLRTVEKKKPTDSLGISIAGGVGSLGADVPIFTAMHPTGVAANO 840

DB 781 sskknmalaseitqglrtvekkkptdslgisiaaggvsgslgadvpiftamhptgvaano 840  
DB 841 KLRVGRDRIYTCSTGEMWHTQAVNLLKNAAGSIEQVAVAGDVSVYTGHHDEPASSL 900  
841 klrvgrdriytcstgemwhtqavnnllknagsgsieqvavagdvsyvtghhdepassl 900  
901 SFTGLTSTIFDQDLGFPPOCKSTLERGPDLGFSIYGVGSPHGDPIYKTVFAKGA 960  
901 sftglststifdqdlgfpockstlergpdlgfsiyvgvgsphgdpiyktvfaakga 960  
961 SEDGRLKRGDQIIIVANGQSLGVTHEBAVALKRTKCTVTLMLVLS 1005  
961 sedgrlkrqdqiivangqslgvtthebavalkrtkctvtlmlvls 1005

RESULT 2

AA04730  
ID AA04730 standard; Protein; 1373 AA.

XX AA04730;

DT 06-JUL-1999 (first entry)

DE Protein containing PDZ domain from clone 38-2-1.

KM PDZ domain; gene expression; human umbilical vascular endothelial cell;

KM HUVEC; stimulation; tumor necrosis factor; TNF; protein binding;

KM cell; proliferation disorder; cancer.

OS Homo sapiens.

PN WO907846-A1.

PD 18-FEB-1999.

PF 12-AUG-1998; 98WO-JP03603.

PR 19-JUN-1998; 98JP-0189944.

PR 12-AUG-1997; 97JP-0230356.

XX (CHUG-) CHUGAI RES INST MOLECULAR MEDICINE INC.

XX Funahashi S, Miyata S;

PI WPI; 1999-167423/14.

DR N-PDB; AAX29908.

PT Protein containing PDZ domain, whose expression is enhanced by TNF

PT stimulation - plays an important role in protein/protein

PT interactions and is used for screening for proteins for use in

PT treatment of cell proliferation disorders such as cancer

XX Claim 1; Page 77-84; 240pp; Japanese.

PS This sequence represents a new protein containing a PDZ domain encoded

CC by the clone 38-2-1, whose expression in human umbilical vascular

CC endothelial cells (HUVEC) is enhanced by stimulation with tumor necrosis

CC factor (TNF) alpha. The new protein is used to identify proteins which

CC bind to it (particularly to the PDZ domains) and the genes encoding them,

CC for use in the treatment of cell proliferation disorders such as cancer.

Sequence 1373 AA:

Query Match 100.0%; Score 5085; DB 20; Length 1373;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1005; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MLQNVSKSEFRTINIAKNSLSGMTVSANKDGLMIVRSIIHGATSRDRIAGPCIL 60  
1 mlgvnskesfertlniakgnsslgmtvsankdglmivrsiihngatstrdriagpcil 60  
61 SINESTISVTNAQARAMLRRHSLIGPDIKITYVPAHELEEFKISLGQSGRNALDIFS 120  
61 sinestisvtnaqaramlrrhsligpdikityvpaheleefkislsgsgrnaldifs 120  
121 SYTRDIPELPEREGESELONTAVSMNORRVELMREPKSICISIVGRGMSR 180  
121 sytrdipelperegeeselnatavsmnorrvellmrepskislsgivgrgmshr 180  
181 LSNGEVMKGIKIKHVELESPAGKNGTLKPCDRIEAPSSQSESEPEKAPLCSVPPPSAF 240  
181 lsngemvngikihvleespagkngtlkpgdrievapsqsepekaplcsvpppsaf 240  
241 AEMGSDHTOSSAKTISQDVKEDDFGYSWKIRERYGTGLGELMIELEKSHSLGSLA 300  
241 aemgsdhtossaktisqdvkeddfgywknirerygtglgelmielekshslgsl 300  
301 GNRKSRMSVITVIGIDPGAAGKDRLOIADELLEINGQILYGRSHONASSITICAPSKV 360  
301 gnrksrmsvityigidpgaagkdrloiadelleingqilygrshonassiticapskv 360  
361 KIIFIRNKDAVNQAVCPGNAVEPLPSNSENLQNKREPTVTSDAAVDLSSFKNVOLLE 420  
361 kifirnkdaavnqavcpгнаveplpsnsenlqnkreptvttsdaavdlssfkvnv 420  
421 LPKDGGGLGIAISEEDTISGVTIKSLTEHGVAAFDGRKVGDDQLAVDEIVGPIEKE 480  
421 lpkdggglgiaiseedtsgvtyiksltehgvaafdgrrkvgddqlavdeivgpiek 480  
481 ISLTKTKMFWKLTITHAENPDSQAVPSAAGASEKKNSGSLWVPOSGSEPSINTS 540  
481 isltkktkmfwkltithaenpdsqavpsaagasekknsgslwvpsgspepsints 540  
541 RSTPAIFASDPATCPIIPGCERTIEISKGTGSLISVGGSDTLGAFIIEVYEEGAA 600  
541 rstpaifasdpatcpiipgcertieiskgtgslisvgsdltlgafiihevyeegaa 600  
601 CKDRLWAGDQIILEVNGIDLRKATHDEALINVLROTPORVRLTLKRDAPRYKEEVCOTLT 660  
601 ckdrlwagdqiilevngidlrrkathdealinvlrtpprvrltlkyrdaprykeevcot 660  
661 IELKKPKGKGLSIVGKRNDDGVFVSDIYVGAADPDGRLIOGDOLLVNGEDVRNASQ 720  
661 ielkkpkkgkglsvgkrrndgvfvsdiyvgaadpdgrrliogdollvngedvrnasq 720  
721 EAVAAALKCSLGTVTEVGRKAPFHSERRPSQTSQVSEGLSSFFPPLSGSSTSESL 780  
721 eavaallkcslgtvtevgrrkapfhserrpsqtsqvseglssffpplsgsstse 780  
781 SSKKNMALASEITQGLRTVEKKKPTDSLGISIAGGVGSLGADVPIFTAMHPTGVAANO 840

```

QY 61 SINEESTISTVNAQARAMLRRHSLIGPDITKITYVPAHLEEFKISIGQSGRMALDIFS 120
   |||||||
Db 429 slineestistvnaqaramlrrhsligpdikityvpaehleefkistigqsgymaldifs 488
QY 121 SYTGNDIPELPEREGEGESELONTAYSNMNPARRVELMREPSKLSIGISVGRMGSR 180
   |||||||
Db 489 sytgndipelperegeeseelqntaysnmnprrvelwrepsklsigisvgrmgssr 548
QY 181 LSNGEVWRGFIKRVHVEDSPAGKNGTLKPGDRIVEAPRSQSESEPEKAPLCSVPPPPSAR 240
   |||||||
Db 549 lsngevwrvgfikhvledspagkngtlkpgdriveaprsqsesepelcsvpppppsar 608
QY 241 AEMGSDHTOSASAKISODVDKEDDFGYSWKNIERYGTLTGELHMIIELEKSHGSLISLA 300
   |||||||
Db 609 aemgsdhtosasaakisdvdkedefgyswknierygtltgelhmiielekshgslisla 668
QY 301 GNKDRSMYSVFIYIDPNGAAGKDGRLQIADLELLEINGOITLYGRSHOMASSIIRKAPSKV 360
   |||||||
Db 669 gnkdrsmysvfiyidpngaagkdgrlqiadelleleingoitlygrshomassilrkapskv 728
QY 361 KIIFIRNKDAVNOMAVCPGNAVEPLPSNSENLONKETEPTVTTSDAAYDLSSFRKNVOHLE 420
   |||||||
Db 729 kifirnkdaavnogmavcpgnaveplpsnsenlnketepvttsdaavdlssfrknvhle 788
QY 421 LPKDGGGLGIAISEEDTLSCVILIKSLTEHGVATDGRILKVGDOILAVDEITVGYPIEKF 480
   |||||||
Db 789 lpkdggglglaiseedtlsgvilksltehgvatdgrilkvqddllavdeitvgyplekf 848
QY 481 ISLLTKAKMTVKLTTHAENPDQAVPSAAGASGEKKNSSQSLMVPDQSGPEPESIRNTS 540
   |||||||
Db 849 islltkakmtvklthhaenpdqavpsaagasekknssqslmvpdsgsperesirnts 908
QY 541 RSTTRPAIFASDPATCPRIPECCETTIEISKRTGLGSIYVGSOTLGCAGFIIEHYEEGAA 600
   |||||||
Db 909 rsttrpaifasdpatcpripeccettieiskrtrlglsiygsotllgcagfiiehyeeгаа 968
QY 601 CKRGRLMAGQOILEVNGIDIRKATHDEAINVLTROTPORVLTLYRDPAPKEEVCDDTLT 660
   |||||||
Db 969 ckrgrlmagqoilevngidirkathdeainvltrotprvltlyrdepapeevecdltl 1028
QY 661 IELQKKRPGKGLSIYVKRNDTGVFVSDIYKGIADPDGRILIOGDOILLVNGEDVNRASQ 720
   |||||||
Db 1029 ielqkkpvgkglsiyvkndtgvfvsdlvkgladpdrilgldqdllvngedvnrnasq 1088
QY 721 EAVVAILLKCSLGVTVLEVEGRIRKAGPRHSERRPSCQTSQVSEGSLSSTFRPLSSGTSSEGLE 760
   |||||||
Db 1089 eavvailkcslgvtvleevgrirkagprhserrpsqtsqvsesslsstfrplssgstsegle 1148
QY 781 SSSKKNALASEIGGLRTPVEKKKGPRTDSIGTSIAGVGSPLGDVPIFTAMMHPTGVAAQTO 840
   |||||||
Db 1149 ssskknalaseigglrtvkekkgprtdsigtsiagvgspdgdvpiftammhptgvaaqtq 1208
QY 841 KLVNNGRIYVIGCTSTEGMHTQAVNLKNAAGSIEQVAVAGDVSVVYTGHHOEPASSSL 900
   |||||||
Db 1209 klvnngriyivigtstegmhtqavnlknagssiemqvavagdvsvvtyghhpepasssl 1268
QY 901 SFMTGLSTSTIFQDDLPGRPOCKSTTLERBPGLGFSIYGVGSPHGDIPITYKTVFAKGA 960
   |||||||
Db 1269 sfmtglststifqddlpgrpocksttlerbpdlgfsiyvgysphgdipityktvfakгаа 1328
QY 961 SEQGRILKRGDOITAVNGOSLEGVTHEEVAVALILKRTGTVLAMLVS 1005
   |||||||
Db 1329 seqgrilkrqgditavngoslegvtheevaivalilkrktgvtlamlvis 1373

```

```

DE Protein containing PDZ domain from clone 38-2-1a.
XX
KW PDZ domain; gene expression; human umbilical vascular endothelial cell;
KW HUVEC; stimulation; tumour necrosis factor; TNF; protein binding;
KW cell; proliferation disorder; cancer.
OS Homo sapiens.
XX
PN WO9907846-A1.
XX
PD 18-FEB-1999.
XX
PF 12-AUG-1998; 98WC-JP03603.
XX
PR 19-JUN-1998; 98JP-0189944.
XX
PR 12-AUG-1997; 97JP-0230356.
XX
PA (CHUG-) CHUGAI RES INST MOLECULAR MEDICINE INC.
XX
PI Funahashi S, Miyata S;
XX
DR WPI: 1999-167423/14.
XX
DR N-PSDB: AAX29908.
XX
PT Protein containing PDZ domain, whose expression is enhanced by TNF
PT stimulation - plays an important role in protein/protein
PT interactions and is used for screening for proteins for use in
PT treatment of cell proliferation disorders such as cancer
XX
PS Claim 1; Page 158-167; 240pp; Japanese.
XX
CC This sequence represents a new protein containing a PDZ domain encoded
CC by the clone 38-2-1a, whose expression in human umbilical vascular
CC endothelial cells (HUVEC) is enhanced by stimulation with tumour necrosis
CC factor (TNF) alpha. The new protein is used to identify proteins which
CC bind to it (particularly to the PDZ domains) and the genes encoding them,
CC for use in the treatment of cell proliferation disorders such as cancer.
XX
SQ Sequence 2000 AA:

Query Match 100.0%; Score 5085; DB 20; Length 2000;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1005; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLOVNSKESEFERTINIAKGNSSLGMTVSANKDGLGIVRSIHGAIISRDGRIAGDCIL 60
   |||||||
Db 996 mlgvnskesefertlnakgnsslgmtvsankdglgmivrsihngaisrdgriaidcill 1055
QY 61 SINEESTISTVNAQARAMLRRHSLIGPDITKITYVPAHLEEFKISIGQSGRMALDIFS 120
   |||||||
Db 1056 slineestistvnaqaramlrrhsligpdikityvpaehleefkistigqsgymaldifs 1115
QY 121 SYTGNDIPELPEREGEGESELONTAYSNMNPARRVELMREPSKLSIGISVGRMGSR 180
   |||||||
Db 1116 sytgndipelperegeeseelqntaysnmnprrvelwrepsklsigisvgrmgssr 1175
QY 181 LSNGEVWRGFIKRVHVEDSPAGKNGTLKPGDRIVEAPRSQSESEPEKAPLCSVPPPPSAR 240
   |||||||
Db 1176 lsngevwrvgfikhvledspagkngtlkpgdriveaprsqsesepelcsvpppppsar 1235
QY 241 AEMGSDHTOSASAKISODVDKEDDFGYSWKNIERYGTLTGELHMIIELEKSHGSLISLA 300
   |||||||
Db 1236 aemgsdhtosasaakisdvdkedefgyswknierygtltgelhmiielekshgslisla 1295
QY 301 GNKDRSMYSVFIYIDPNGAAGKDGRLQIADLELLEINGOITLYGRSHOMASSIIRKAPSKV 360
   |||||||
Db 1296 gnkdrsmysvfiyidpngaagkdgrlqiadelleleingoitlygrshomassilrkapskv 1355
QY 361 KIIFIRNKDAVNOMAVCPGNAVEPLPSNSENLONKETEPTVTTSDAAYDLSSFRKNVOHLE 420
   |||||||
Db 1356 kifirnkdaavnogmavcpgnaveplpsnsenlnketepvttsdaavdlssfrknvhle 1415

```

```

OY 421 LPKDGGGLGIAISEEDTLGVIKSLTEHGVANATDGLKVGDOILLAVDEIVVGIPIKEF 480
DB 1416 lpkdggglgialseedtlsgvllkeltehgvaaatdgrllkvvgdillavddelivvgpliekf 1475
OY 481 ISLTLTAKMTVTLTHAENPDQAAPSAGAAAGSEKKNSQSGLAMPQSGSPPESEIRMTS 540
DB 1476 ISLTLTAKMTVTLTHAENPDQAAPSAGAAAGSEKKNSQSGLAMPQSGSPPESEIRMTS 1535
OY 541 RSRTPAIFASDPATCPPIPGCETTEIEISKGRGTLGSLIVGSDTLTGAFIHEVEEGAA 600
DB 1536 rsrtpaifasdpacpplpgcetteliskgrtglsglsvgsdltlgafilheveyegaa 1595
OY 601 CKDGLMAGDOILEVNGIDLRKATHDEAINVLKQTPQVRVRLTYRDEAPYKEEVCDFLT 660
DB 1596 ckdgllmaggdillevngidlrkathdeainvlrqpvrvtllyrdeapykeevcdfilt 1655
OY 661 IELOKKPKGKGLSTVGKRNQTVGVSDIVKGINADPQGRILQSGOILLVNGEDVRNNSQ 720
DB 1656 IELGKKPKKGLSTVGKRNQTVGVSDIVKGINADPQGRILQSGOILLVNGEDVRNNSQ 1715
OY 721 EAVALLKCSLQTVTLVLEVGRIKAGPFHSESRPSQTSQVSEGLSFTPLSGSSTSESL 780
DB 1716 EAVALLKCSLQTVTLVLEVGRIKAGPFHSESRPSQTSQVSEGLSFTPLSGSSTSESL 1775
OY 781 SSSKKNALASERIGLRTVEMKKGPDSLCISLAGCVSPLGDVPFIAMHPTGVAAQTQ 840
DB 1776 SSSKKNALASERIGLRTVEMKKGPDSLCISLAGCVSPLGDVPFIAMHPTGVAAQTQ 1835
OY 841 KLRVGDRIYTCGTSTEGMTHQAANLKNASGSTEMOVVAGGVSVYTGHHQEPASSSL 900
DB 1836 KLRVGDRIYTCGTSTEGMTHQAANLKNASGSTEMOVVAGGVSVYTGHHQEPASSSL 1895
OY 901 SFTGLTSTSIPODLGPPQCKSTILKRPDGLGFSIVGSGPHGDLPIYKTVPAKGA 960
DB 1896 sftglststsiPODLGPPQCKSTILKRPDGLGFSIVGSGPHGDLPIYKTVPAKGA 1955
OY 961 SEDGLKRGDOIAVNGOSLEGVTHEAVAILTKRKGTVTLMVLS 1005
DB 1956 sedglkrgdOIAVNGOSLEGVTHEAVAILTKRKGTVTLMVLS 2000

```

RESULT 4  
 ID AAY04733 standard; Protein: 2070 AA.  
 AC AAY04733;  
 DT 06-JUL-1999 (first entry)  
 DE Protein containing PDZ domain from clone 38-2-1b.  
 KW PDZ domain; gene expression; human umbilical vascular endothelial cell; HUVEC; stimulation; tumour necrosis factor; TNF; protein binding;  
 KW cell; proliferation disorder; cancer.  
 OS Homo sapiens.  
 PN MO9907846-A1.  
 PD 18-FEB-1999.  
 PF 12-AUG-1998; 98WO-JP03603.  
 PR 19-JUN-1998; 98JP-0189944.  
 PR 12-AUG-1997; 97JP-0230356.  
 PA (CHUG-) CHUGAI RES INST MOLECULAR MEDICINE INC.  
 PI Funahashi S, Miyata S;  
 DR MPI: 1999-167423/14.  
 DR N-PSDB: AAX29910.  
 XX

PT Protein containing PDZ domain, whose expression is enhanced by TNF  
 PT stimulation - plays an important role in protein/protein  
 PT interactions and is used for screening for proteins for use in  
 PT treatment of cell proliferation disorders such as cancer

PS Claim 1; Page 167-176; 240pp; Japanese.

CC This sequence represents a new protein containing a PDZ domain encoded  
 CC by the clone 38-2-1b, whose expression in human umbilical vascular  
 CC endothelial cells (HUVEC) is enhanced by stimulation with tumour necrosis  
 CC factor (TNF) alpha. The new protein is used to identify proteins which  
 CC bind to it (particularly to the PDZ domain) and the genes encoding them,  
 CC for use in the treatment of cell proliferation disorders such as cancer.

XX Sequence 2070 AA:

SO

Query Match 99.0%; Score 5035; DB 20; Length 2070;  
 Best Local Similarity 93.4%; Pred. No. 0;  
 Matches 1004; Conservative 0; Mismatches 1; Indels 70; Gaps 1;

```

OY 1 MLOANSKSEFRTINIAKNSLIGMTVANRDKGIMYRSTIHGATSRDRIRAGCIL 60
DB 996 mlgnsksesfertlniaknsligmtvanrkdglmyrslthgatastdriragdcil 1055
OY 61 SINESTSVTNAQARAMLRRHSLIGPDIKITYPAEHLFEFKISLGQSGRVNADIFS 120
DB 1056 slnestsvtnaqaramlrrhsligpdikitypaehlfeekslsgqsgrvnadtifs 1115
OY 121 SYTGRIPELPEREGEREESELDONTATSNMORRVVLRKPEPSLGSIVGRGMSR 180
DB 1116 sytgripeleperegereseldontatSNMORRVVLRKPEPSLGSIVGRGMSR 1175
OY 181 LSNCEVMGIRFIKHVLESPAGKNGTLKPCGRIVE----- 215
DB 1176 lsncevmgirikhvleSPAGKNGTLKPCGRIVE----- 1235
OY 216 -----APSQSESEPKAPLC 230
DB 1236 vfmwgsllnrprkspjpsllhnlpykynfssstnpfadsldqnadapagsesepkajlc 1295
OY 231 SVPPPPSAFAPEMGSDHROSSASKRISODVDEKEDERGYSKNIIRERYGTLNGLMIELEK 290
DB 1296 svppppsaafapemgSDHROSSASKRISQDVDEKEDERGYSKNIIRERYGTLNGLMIELEK 1355
OY 291 GHSGLGLSLAKNDRSRMSVFIYGDIPNGAAGKGRQLQIADLELLEINQOILYGRSHQNAS 350
DB 1356 ghsghlglslakndrSRMSVFIYGDIPNGAAGKGRQLQIADLELLEINQOILYGRSHQNAS 1415
OY 351 SIICAPSKVKIIFIRNKDAVNQMAVCPGNAVEPLPSNSENLOKKEPEPTVTSDAAYDL 410
DB 1416 sliicapskvkiifirnkDAVNQMAVCPGNAVEPLPSNSENLOKKEPEPTVTSdaaydl 1475
OY 411 SSKFNVQHLLEPKOOGGLGIAISEEDTLGVIKSLTEHGVANATDGLKVGDOILLAVDE 470
DB 1476 sskfnvqhllEPKOGGLGIAISEEDTLGVIKSLTEHGVANATDGLKVGDOILLAVDE 1535
OY 471 IVVGYPIEKFTSLTAKMTVTLTHAENPDQAAPSAGAAAGSEKKNSQSGLAMPQSGS 530
DB 1536 ivvgypiekftslTAKMTVTLTHAENPDQAAPSAGAAAGSEKKNSQSGLAMPQSGS 1595
OY 531 PEPESIRNTSRSSTPAIFASDPATCPPIPGCETTEIEISKGRGTLGSLIVGSDTLTGAFI 590
DB 1596 pepesirntsrssTPAIFASDPATCPPIPGCETTEIEISKGRGTLGSLIVGSDTLTGAFI 1655
OY 591 IHEVYEEGAACKDGRMAGDOILEVNGIDLRKATHDEAINVLKQTPQVRVRLTYRDEAPY 650
DB 1656 IheveyeegaackdgrMAGDOILEVNGIDLRKATHDEAINVLKQTPQVRVRLTYRdeapy 1715
OY 651 KEEVCDFLTLELOKKPKGKGLSTVGKRNQTVGVSDIVKGINADPQGRILQSGOILLV 710
DB 1716 keevcdfiltlelGKKPKGKGLSTVGKRNQTVGVSDIVKGINADPQGRILQSGOILLV 1775

```



QY 711 NGEVNNASQEAVALLKSLGTVLEVGRIKAGPFHSRRPSQTSQVSEGLSFTFPL 770  
| | | | |  
Db 1776 ngevdvnaaqaavallkslgtvlevgrikagphserpsqtsqveglstftpl 1835  
| | | | |  
QY 771 SSSSTSESSSSKKNALASEIOGLFTVMKKGPDSLSIAGVGSPLGDVPIFIAM 830  
| | | | |  
Db 1836 sgsstseesssskknalaseloqlrtvemkkypdtslgsiagvgsplgdvplfiann 1895  
| | | | |  
QY 831 HPTGVAAGOKIRVGRDRIYTGISTEGTHTHOAVALLKNASSTIMOVAVAGDVSVTG 890  
| | | | |  
Db 1896 hptcvaagqkrlrvgrlvtlqgstegmthcqaullknasgslemqvvaagdvsvvtg 1955  
| | | | |  
QY 891 HHQEPASSLSFTGLSTSFODDLPPOCKSTTERGPDGLGFSIVGSGSPHCDLPY 950  
| | | | |  
Db 1956 hhqepasslsftglststlfgddlgppqckstlergpdglgfsivgsgsphcdlpy 2015  
| | | | |  
QY 951 VKTVEAKGAASEDGRLKRGDQIIIVANGSLGEGVTHEEAVAILKRTKGTVLAWLS 1005  
| | | | |  
Db 2016 vktvfakgaasedgrlkrqdgqllavngsllegvtheeavaailkrtkgtvtlamlvs 2070  
| | | | |

## RESULT 5

AAV53753 standard; Protein; 2037 AA.

AAV53753;

22-FEB-2000 (first entry)

Amino acid sequence of the MMS2 protein.

Human: MMS2; MMAC1; PDZ domain; tumour suppressor; tyrosine phosphatase;  
scatolding protein; cancer.

Homo sapiens.

WO958548-A1.

18-NOV-1999.

07-MAY-1999; 99WO-US09969.

08-MAY-1998; 98US-0084740.

(MYRI-) MYRIAD GENETICS INC.

Bartel PL, Tavtigian SV;

WPI; 2000-053077/04.

N-PSDB; AA236453.

Nucleic acids and polypeptides representing human MMS2, useful for  
detecting, diagnosing a predisposition to, and treating cancer -

Claim 1; Page 93-99; 112pp; English.

The present sequence represents human MMS2 protein. The MMAC1 protein binds to MMS2. The MMS2 protein has 11 post-synaptic density protein, disc-large, zc-1 (PDZ) domains and one or more of these domains interact specifically with the carboxyl terminal amino acids of MMAC1 (see AAV53754). Specifically, it appears that domain 7, 10 and 13 interact with MMAC1. Since MMS2 contains 11 PDZ domains and interacts with MMAC1, a known tumour suppressor having a region of homology with protein tyrosine phosphatases, MMS2 acts as a scatolding protein in a common biological pathway with MMAC1. It is believed that the interaction between MMAC1 and MMS2 is required for the tumour suppressor activity of MMAC1. The MMS2 polypeptides, polynucleotides, fragments and specific or complex specific antibodies may be used for detecting cancer or a predisposition to cancer and screening for agents that may be used to treat MMS2 and/or MMAC1 related cancer. The polypeptides and polynucleotides may also be used to treat cancer.

Sequence 2037 AA;

Query Match 98.7%; Score 5018.5; DB 21; Length 2037;  
Best Local Similarity 95.7%; Pred. No. 0;  
Matches 997; Conservative 5; Mismatches 3; Indels 37; Gaps 1;

QY 1 MLOVNSKEEFERTINAKNSSIGMTVSANKDGLIVRSIIHGGAISDGRIGADICIL 60  
| | | | |  
Db 996 mlgvnskeferrtinaknsslgtmtvsankdglivrsiilhggaistdgrilaigdcil 1055  
| | | | |  
QY 61 SINEESTISVTNAQARAKMRRLSLGPDITIKITYPAEHLKEFRISLGQSGRYMALDIFS 120  
| | | | |  
Db 1056 sineestisvtnaagaramlrmlslgpditkitypaeheleekfislsgqsgrymalldifs 1115  
| | | | |  
QY 121 STTGRIPELPEREBSGESELONTAYSNMNPRAVELMREBPSLSGISYGRKMGSR 180  
| | | | |  
Db 1116 sttgripeleperesgeeselontaysnmnprravelmrebpslsigisygrmgshr 1175  
| | | | |  
QY 181 LSNGEVNRGICFIKHYLEDSPACKNGTLKRGDRIVE----- 215  
| | | | |  
Db 1176 lsngevnmrgiflkhvledspackngtlkpgdrilvevdgmdlrdashegavealrkagpvr 1235  
| | | | |  
QY 216 -----APQSSESEPEKAPLCVPPPPSAFAEMGSDHNGSASKISQVDXED 263  
| | | | |  
Db 1236 vfmvgsilnrprapsgsesepckaplcsvppppsafeemgshgssasklsqdvdxed 1295  
| | | | |  
QY 264 EFGYSWKNIIRERYGTLTGELHMIIELEKSHGLSLAGLAKKDRSRMSVFTVGDIPNGAOK 323  
| | | | |  
Db 1296 efgyswkniirerygtltgelhmiielekshgslaglakkdrrmsvftvgdipngaaok 1355  
| | | | |  
QY 324 DGRLOIADLELEINQOILKGRSHONASTIICAPSVKIIIFIRNKAVNMVCPENAYE 383  
| | | | |  
Db 1356 dgrloiadelleinqoilkgrshonastiiicapsvkiiifirnkavnmvcpenaave 1415  
| | | | |  
QY 384 PLPSNSENLOKNETEPTVTTSDAAYDLSSFKKNVHLEPKRDQGLGIAISEEDTLGVI 443  
| | | | |  
Db 1416 plpsnseulqnketepvtvtsdaavolsfknvqhlepkdggjglaiseedtlsgvll 1475  
| | | | |  
QY 444 KSLTEHGAATDGRKLVGQOIIAVIDEIVGYPIEKFISLTKAKTFVLTTHAENPDSO 503  
| | | | |  
Db 1476 ksltehgaatdgrlkvgqoiiavideivgyplekfisltkaktvltlhaenpdsq 1535  
| | | | |  
QY 504 AVPSAAGAASGERKNSOSLAMPQSGPEPESIRNRSSTPAIFASDPATCPIIGCET 563  
| | | | |  
Db 1536 avpsaagaasgerknsqslampqsgpepestirnrstspafasdpatcpiigcet 1595  
| | | | |  
QY 564 TIEISKRTGLSLIVGSDTLGAFIIEHYVEGAACKDGRUMAGDQIILEVNGIDLKRA 623  
| | | | |  
Db 1596 tieiskrtglslivgsdtilgafiihyveegaackdgrlwgddqlllevngidlrlra 1655  
| | | | |  
QY 624 THDEAINVRQTPORVRLTLRDEAPYKEEVCDDTLTELQKPKGGLSLYGRKNDPG 683  
| | | | |  
Db 1656 thdeainvrltpqrvrltlrdeapykeevcdtltelqkpkgyglslvgrkndpg 1715  
| | | | |  
QY 684 VEVSDIVKGCIGADPDGRLIQGDQIILVNGEDVYRNASQEAVALLKSLGTVLEVGRIKA 743  
| | | | |  
Db 1716 vevsdilvkgigadaadgrlimgddqilnvngedvyrnaqaavallkslgtvlevgrika 1775  
| | | | |  
QY 744 GFPHSERRPSQTSQVSEGLSFTFPLSGSSTSESSSSKKNALASEIOGLRTVEMKKG 803  
| | | | |  
Db 1776 gpfhserrpsqtsqveglstftplsgsstseesssskknalaseloqlrtvemkkgy 1835  
| | | | |  
QY 804 PDSLSLSIAGVGSPLGDVPIFIAMHPTGVAAGQOKIRVGRDRIYTGISTEGTHTHQ 863  
| | | | |  
Db 1836 pdslsislagvgsplgdvplfiammhptcvaagqkrlrvgrlvtlqgstegmthcqa 1895  
| | | | |  
QY 864 AVNLKKNAGSITEMQVAVAGDVSVTGHHQEPASSLSFTGLSTSFODDLPPOCKSKI 923  
| | | | |  
Db 1896 avnlknasgsitemqvavagdvsvvtgghqepasslsftglststlfgddlgppqcksl 1955  
| | | | |  
QY 924 TLERGPDGLGFSIVGSGPHCDLPYKTVFAKGAASEDGRLKRGDQIIIVANGSLGEGV 983  
| | | | |  
Db 1956 tlergpdglgfsivgsgphcdlpykvtvfkagaasedgrlkrqdgqllavngsllegv 2015  
| | | | |

OY 984 THEEAVALIKRTKGTVTLMVLIS 1005  
 |||||  
 Db 2016 theeaavallkrtkgtvtlmvlis 2037

RESULT 6  
 AAB01383  
 ID AAB01383 standard; Protein: 856 AA.  
 XX  
 AC AAB01383;  
 XX 20-OCT-2000 (first entry)  
 XX  
 DE Neuron-associated protein.

XX Neuron associated protein: NEUP; neurological disorder; epilepsy;  
 KW Ischemic cerebrovascular disease; stroke; cerebral neoplasm;  
 KW Alzheimer's disease; Pick's disease; Huntington's disease;  
 KW dementia; Parkinson's disease; demyelinating disease; meningitis;  
 KW prion disease; Kuru; Creutzfeldt-Jakob disease; neurofibromatosis;  
 KW cerebral palsy; muscular dystrophy; central nervous system; CNS;  
 KW peripheral nervous system; PNS; myopathy; schizophrenia;  
 KW actinic keratosis; arteriosclerosis; atherosclerosis; buritis;  
 KW cirrhosis; hepatitis; mixed connective tissue disease; MCTD;  
 KW myelofibrosis; paroxysmal nocturnal haemoglobinuria; cancer;  
 KW autoimmune disease; inflammation; acquired immunodeficiency syndrome;  
 KW AIDS; Addison's disease; adult respiratory distress syndrome;  
 KW allergy; ankylosing spondylitis; amyloidosis; anaemia; asthma;  
 KW Werner syndrome; trauma; human.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Domain 1..65  
 FT Modified-site 18 /label- PDZ domain signature  
 FT Modified-site 20 /note- "Potential glycosylation site"  
 FT Modified-site 69 /note- "Potential phosphorylation site"  
 FT Modified-site 71 /note- "Potential phosphorylation site"  
 FT Modified-site 73 /note- "Potential phosphorylation site"  
 FT Modified-site 118 /note- "Potential phosphorylation site"  
 FT Modified-site 119 /note- "Potential phosphorylation site"  
 FT Modified-site 136..218 /note- "Potential phosphorylation site"  
 FT Domain 144..147  
 FT Modified-site 199 /note- "Glycosaminoglycan attachment site"  
 FT Modified-site 238 /note- "Potential glycosylation site"  
 FT Modified-site 253 /note- "Potential phosphorylation site"  
 FT Modified-site 263 /note- "Potential phosphorylation site"  
 FT Modified-site 269..349 /note- "Potential phosphorylation site"  
 FT Domain 284 /label- PDZ domain signature  
 FT Modified-site 296 /note- "Potential phosphorylation site"  
 FT Modified-site 337 /note- "Potential phosphorylation site"  
 FT Modified-site 341 /note- "Potential phosphorylation site"  
 FT Modified-site 369 /note- "Potential phosphorylation site"

FT Modified-site 370 /note- "Potential glycosylation site"  
 FT Modified-site 386 /note- "Potential phosphorylation site"  
 FT Modified-site 389 /note- "Potential phosphorylation site"  
 FT Modified-site 390 /note- "Potential glycosylation site"  
 FT Modified-site 414 /note- "Potential phosphorylation site"  
 FT Modified-site 415..497 /note- "Potential phosphorylation site"  
 FT Domain 475 /label- PDZ domain signature  
 FT Modified-site 511..592 /note- "Potential phosphorylation site"  
 FT Domain 531 /label- PDZ domain signature  
 FT Modified-site 568 /note- "Potential glycosylation site"  
 FT Modified-site 599 /note- "Potential glycosylation site"  
 FT Modified-site 604 /note- "Potential phosphorylation site"  
 FT Modified-site 625 /note- "Potential phosphorylation site"  
 FT Modified-site 633 /note- "Potential phosphorylation site"  
 FT Modified-site 634 /note- "Potential phosphorylation site"  
 FT Modified-site 648..733 /note- "Potential phosphorylation site"  
 FT Domain 690 /label- PDZ domain signature  
 FT Modified-site 705 /note- "Potential phosphorylation site"  
 FT Modified-site 721 /note- "Potential phosphorylation site"  
 FT Domain 773..856 /note- "Potential glycosylation site"  
 FT Modified-site 819..821 /label- PDZ domain signature  
 FT Region 835 /label- Cell attachment sequence  
 FT Modified-site 835 /note- "Potential phosphorylation site"

WO200034477-A2.

15-JUN-2000.

10-DEC-1999; 99WO-US30408.

11-DEC-1998; 98US-0210083.

11-DEC-1998; 98US-9123456.

09-FEB-1999; 99US-0119365.

16-MAR-1999; 99US-0124687.

(INCY-) INCYTE PHARM INC.

Tang YT, Yue H, Baughn MR, Hillman JL, Lai P, Au-Young J, Yang J;

Lu DAM, Azimzal Y;

WPI: 2000-423423/36.

N-PSDB; AAA47424.

New human neuron-associated proteins and polynucleotides encoding them,  
 useful for diagnosis, treatment and prevention of cell proliferative  
 disorders including cancer, neuronal and neurological disorders  
 Claim 1; Page 105-108; 145pp; English.  
 Human neuron-associated proteins (NEUP) can be used for  
 treating or preventing a disorder associated with decreased



Query Match 75.5% Score 3840; DB 20; Length 763;  
 Best Local Similarity 99.7%; Pred. No. 0;  
 Matches 761; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

243 MGSHTQSSAKISQDDKEDFEFGYKNTFRERYGTLTGELHMLEKGGSLSLAGN 302  
 |||  
 1 mgshtqssakisdvdckedefgyswnfrerygtltgelhmlleghslslagn 60

303 KDRSRMSVFYIGIDPNCAGKGRQIADLELEINGOILYGRSHQNASSTIKCAPSKYKI 362  
 |||  
 61 kdrsrmsvflvgidpncagkgdqrlqldelldengdillygrshqnasstikcapskvki 120

363 IFINKDAVNOMAVCPGNNAVEPLPSNSEFLONKTEPTVTTSDAAYDLSSFKNOHLELP 422  
 |||  
 121 ifinkdavnmavcpgnnavelpnsenflonkteptvttsdaaydlssfknhlelp 180

423 KDQGLGIAISEEDTLGGVITIKSLTEHGVAATDGLRVGQDIIAVDDEIVVGYIEKFI 482  
 |||  
 181 kdqglgiaiseedtlggvilksltehgvaatdglrvvgdiiavddeivvgyplekfls 240

483 LKTKAKMTVKLTIAENPDQAVPSAAGAAAGKKNSSQSILMPQSGPEPESTIRNTSRS 542  
 |||  
 241 lkakmtvkltiaenpdqavpsaagaaagkknssqslmvpqsgpepestirntsrs 300

543 STRAIFASDPATCPPIFGCETTIEISKRTGLGUSIVGSDTLGAFIHEVYEGAAK 602  
 |||  
 301 strafasdpatcpifgcettieiskrtglgusivgsdtllgafihveyeaaack 360

603 DGRIMADQILEVNGIDLRKATHDEATNVLKQTPQVRRLTIYREARYKKEEVCDTLIE 662  
 |||  
 361 dgrimadqilevngidlratkathdeatnvlkqtpqvrlltiyrearykkeevcdtlie 420

663 LQKPRGKGLSTIVKBNQDGVFVSDIVKGIADPDLGDDQILLYNGEDVNNASQEA 722  
 |||  
 421 lqkprgkglstivkbnqdvfvfsvdlvkgiadpdlrgddqillyngedvnnasgea 480

723 VAALKCSLGTIVLEVGRIKAGPRHSERRPSQISVBSGLSFTFPLSGSSSESLESS 782  
 |||  
 481 vaalkcslgtivlevgrikagprhserrpsqisvbsglstfplsgsssesless 540

783 SKKVALSEIOGLRTVMKKGPTDLSLGISTAGVSPGLDPRFIAMMHPGVAAOQKL 842  
 |||  
 541 skkvalseioqlrtvmkkgptdlslgistagvspgl DPRFIAMMHPGVAAQCKL 600

843 RVGDRIYTCSTEGMTHTQAVNLKNASGSIEMQVAVAGDVSVVGHNOEPASSLSLF 902  
 |||  
 601 rvgdriytcstegmthtqavnlknasgsiemqvavagdvsvvghnqepasslsf 660

903 TGLTSTIFQDDDLGPPQCKSTTLERKPDGLGFSIVGYSFHDLPITYKTVFAKGAASE 962  
 |||  
 661 tglststifqdddlgppqcksttlterkpdglgfsivgysfhdlpitykvtvfkgaase 720

963 DGRILRGDQIIAVNGOSLEGVTHEEVAAILKRTKGTWTLWLS 1005  
 |||  
 721 dgrilrgdqiiavngoslegvtheevaailkrtkgtwtlwls 763

RESULT 8.  
 AAY24025  
 ID AAY24025 standard; Protein; 1881 AA.  
 AC AAY24025;  
 XX 29-SEP-1999 (first entry)  
 DE Amino acid sequence of the human MMS1 protein.  
 KW Human; MMS1 protein; MMS1 interacting protein; tumour suppression;  
 KM MMAC1 pathway; immunogen; cancer; cell neoplastic growth.  
 OS Homo sapiens.  
 XX  
 PN W0936566-A1.

XX 22-JUL-1999.  
 PD  
 XX 19-JAN-1999; 99MO-US00995.  
 PF  
 XX 20-JAN-1998; 98US-0071861.  
 PR  
 XX (MYRI-) MYRIAD GENETICS INC.  
 PA  
 XX Bartel PL, Tavtigian SV;  
 PI  
 XX WPI; 1999-458472/38.  
 DR  
 XX N-PSDB; AAX86366.  
 PT  
 PT MMS1, an MMAC1 (tumour suppressor) interacting protein and related  
 PT polynucleotides  
 PS  
 XX Claim 14; Page 88-93; 107pp; English.

CC The present sequence represents a MMS1 protein. The protein is a MMAC1  
 CC interacting protein which is involved in tumour suppression activity  
 CC in the MMAC1 pathway. MMS1, antigenic fragments or fusion proteins of  
 CC these are used as immunogens for antibody production. Primers derived  
 CC from MMS1 genomic clones can be used for identification of MMS1 genes  
 CC and for synthesis by amplification of MMS1 DNA or RNA. Detecting an  
 CC alteration in MMS1 can be used to diagnose cancer. A germline alteration  
 CC in an MMS1 gene is indicative of a predisposition to cancer. A somatic  
 CC mutation in an MMS1 gene is indicative that the tissue is cancerous.  
 CC Analysis of MMAC1 and MMS1 (or PDZ domain 6 of MMS1) binding  
 CC interactions can be used for detection of alterations in MMAC1  
 CC associated with cancer. Wild-type MMS1 or a homologue can be used to  
 CC supply wild-type MMS1 gene function (or a substantially similar  
 CC function) to a cell, which has lost the gene function due to the MMS1  
 CC gene mutation. The gene suppresses neoplastic growth of the cell.  
 CC Transgenic animals having an altered MMS1 can be used as a model for  
 CC identifying drug candidates useful in treating cancer.  
 CC  
 XX  
 SQ Sequence 1881 AA;

Query Match 42.0% Score 2133.5; DB 20; Length 1881;  
 Best Local Similarity 46.5%; Pred. No. 9.1e-170;  
 Matches 499; Conservative 140; Mismatches 229; Indels 205; Gaps 26;

3 QNWSKESFER---TINANGNSISLGMTVSANKDGLMIVRSITHGGAISRGRHAIAGDC 58  
 |||  
 943 enwkenfmeslpsystegnsgqr--fdlenlnslaktsldlmpnd----- 992

59 ILSINESTISVTNAQRAAMLRHSLIGPDIKITY-VPAHELEEFKISLGO-QSGRMAL 116  
 |||  
 993 -----vgpsllldlpvvaqregedilplyghatrlvsk 1027

117 DIFSSYTG-----RDIPELPEREEGEGESLQNTAVSWMNPARRVELMREPSKSLGI 169  
 |||  
 1028 --asaytgmalsryatctceiperegegeec---pnishwpprtivelftepnvalgl 1081

170 SIVGRCMGSRSLNSGEVNRGIFIKHVLDSPPACNGKTLKQDRIVERA-----PSQSESEP 224  
 |||  
 1082 slvggqvltrklngelkglfllkqvledspagfnaclkcgdkllveysdqlqnashea 1141

225 EKA-----PLCSVPPPPPSAF---AEMGSDHQSSASK----- 254  
 |||  
 1142 vealknagnpvflvgsllstprvlpvnhkankltanqndtqekkrqgtappmk1 1201

255 -----ISODVDR-EDFEFGYSWKNIRERYGTLTGELHMLEKGGSLSLAGNKDRSR 307  
 |||  
 1202 pppykaltddseneedafldqrlqldelldengdillygrshqnasstikcapskvki 1261

308 MSYFVIGIDPNCAGKGRQIADLELEINGOILYGRSHQNASSTIKCAPSKYKIPIRN 367  
 |||  
 1262 msflvglngpdaadgrmhlgdelldengdillygrshqnasstikcapskvklvflrn 1321

368 KDAVNOMAVCPGNNAVEPLPSNSEFLONKTEPTVTTSDAAYDLSSFKNOHLELPKQGG 427

```

Db 1322 edavngmavtp-----fypssasp-----sdl-----edgsg 1348
Oy 428 LGAISEEDTLISGVIIKSLTEHGVATDGRKLVGQDIIAVDEIIVGPIEKFISLKT 487
Db 1349 teplasee-----gsllevglkqlpesefkila-----vsgmkqg 1383
Oy 488 KMTVTLTHAENPDQAVPSAAGASGEKKNSOSLWMPQSSPEPESTRNRSSTPAI 547
Db 1384 kypkvwstf-----ssqelplapass-----yhsldadftgyggfagaplsav----- 1423
Oy 548 FASDPATPIIFGCEETTEISKRTGLSIYGSGLTLLGAPTHIEVEEAGAACDGRIM 607
Db 1424 ----pactcpivpggemlleltskgrsglglstivgkdpdnalvheveegsaadgrlw 1480
Oy 608 AGDQILEVNGIDLKRAKHDEAINVLRQTPQRVRLTLVREADYKEEVCDTITLQKKP 667
Db 1481 agdqgllevngvdlrnsheeaialtrqpqrlyvrydeahyrdleenlelfpvdlqka 1540
Oy 668 GKGGLSLVIGKRNQGVFVSDIVKGIADPDGRLIOGDQIILVNGEDVNRNASQEAVAL 727
Db 1541 gfglglstivgkngsvlstdivkgsaadlgrlilgqdgllsvngedmnaqetvatll 1600
Oy 728 KCSLGTVTLVLENGRIKAGFPFHSERPSQTSQVSEGLSFTFP-----LSGSSTESL 779
Db 1601 kcaqglvqlelgirlragvstsaarltsgnsgsgsahsahsahsfapvltglnlvgtkrv 1660
Oy 780 ESSSKNMLASIEIGLRVYEMKKGPTDSIGISAGVSPCLDVLFIAMHPTVAAOT 839
Db 1661 sdpskngs-gldmep-rivelnrelsdalglslagrgspglpdlvfilamqasvvaar 1718
Oy 840 OKLRAGDRIVTICGTSTEGMTHQAVNLKKNAGSITEMOVVGVGVSVY-----TGH 891
Db 1719 qkllvgdrlvslngqpldgishadvnllkngayrlllgvavdthlsaaqlennstgy 1778
Oy 892 HQEPASSLSFTGLTSTIFODLGPPOCKSITLERGPDGLGFSIVGSGSPHGDLPYV 951
Db 1779 h-----lgsptlaehpedteppkiltleksgslgfsivgsgsphgdplpyv 1828
Oy 952 KVFPAKGAASEGRKLRGQDIIAVNGQSLEGTHEEAVAILKRTGTITLAWL 1004
Db 1829 klvfkgaadagrlkrqgdlavngellegvtheqavallkhqrgvclvli 1881

RESULT 9
AAV74150
ID AAV74150 standard; Protein: 318 AA.
XX
AC AAV74150;
XX
DT 14-MAR-2000 (first entry)
XX
DE Human prostate tumor EST fragment derived protein #337.
XX
KW Pancreas; tumor; EST; expressed sequence tag; human; cytostatic;
XX
OS treatment.
XX
XX Homo sapiens.
XX
PN DE19820190-A1.
XX
PD 04-NOV-1999.
XX
PF 28-APR-1998; 98DE-1020190.
XX
PR 28-APR-1998; 98DE-1020190.
XX
PA (META-) METAGEN GES GENOMFORSCHUNG MBH.
XX
PI Rosenthal A, Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E;
XX
DR WPI; 1999-621386/54.
XX
N-PSDB; AA252969.

```

```

XX
PT New human nucleic acid sequences from pancreatic tumors, and related
PT proteins
PS Claim 23; Page 449-450; 502pp; German.
XX
CC This invention describes novel polypeptides and their encoding nucleic
CC acids derived from human pancreatic tumor tissue which have cytostatic
CC activity. The sequences are also useful in producing pharmaceutical
CC compositions for treatment of pancreatic tumors. AAV73814-Y74252
CC represent protein fragments encoded by the human pancreatic tumor cDNA
CC library derived expressed sequence tag (EST) sequences represented in
CC AA52858-453014.
XX
SO .Sequence 318 AA:

Query Match 29.5%; Score 1501; DB 20; Length 318;
Best Local Similarity 96.7%; Pred. No. 7,8e-118;
Matches 297; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

Oy 422 PKDGLGIAISEEDTLISGVIIKSLTEHGVATDGRKLVGQDIIAVDEIIVGPIEKF 481
Db 9 prrgglglalseddlsgvlklslehgvaatdgrllvgdqlavddelivvgylekfl 68
Oy 482 SILTKAKMTVTLTHAENPDQAVPSAAGASGEKKNSOSLWMPQSSPEPESTRNRS 541
Db 69 siltkakmtvklthnaenpdqavpsaagaasgekknsqslmwpqsspepestrnrs 128
Oy 542 SSTPAIFASDPATPIIFGCEETTEISKRTGLSIYGSGLTLLGAPTHIEVEEAGAAC 601
Db 129 sstplfssdpactcpilpgcettelstskgrtglstivgsedtlgalllhevyeegaac 188
Oy 602 KDGRLMAGDQILEVNGIDLKRAKHDEAINVLRQTPQRVRLTLVREADYKEEVCDTIT 661
Db 189 kdgrlmagdqllenvngldlrkathaeainvrlrqpqrlyltyrdeaykeevcdtltl 248
Oy 662 ELQKPRGGLSLVIGKRNQGVFVSDIVKGIADPDGRLIOGDQIILVNGEDVNRNASQ 721
Db 249 elqkprgglslstivgkngsvlstdivkgsaadlgrlilgqdgllsvngedmnaqetvat 308
Oy 722 AVALALK 728
Db 309 avawwllk 315

RESULT 10
AAV04738
ID AAV04738 standard; Protein: 251 AA.
XX
AC AAV04738;
XX
DT 06-JUL-1999 (first entry)
XX
DE PDZ domain-containing protein gene encoded by clone FH750.
XX
KW HUVEC; stimulation; tumor necrosis factor; TNF; protein binding; PCR;
XX
OS cell; proliferation disorder; cancer; primer; amplification.
XX
XX Homo sapiens.
XX
PN WO9907846-A1.
XX
PD 18-FEB-1999.
XX
PF 12-AUG-1998; 98WO-JP03603.
XX
PR 19-JUN-1998; 98JP-0189944.
XX
PR 12-AUG-1997; 97JP-0230356.
XX
PA (CHUG-) CHUGAI RES INST MOLECULAR MEDICINE INC.
XX

```

PI Funahashi S, Miyata S;  
 XX WPI: 1999-167423/14.  
 DR N-PSDB; AAX29969.  
 XX  
 PT Protein containing PDZ domain, whose expression is enhanced by TNF  
 stimulation - plays an important role in protein/protein  
 PT interactions and is used for screening for proteins for use in  
 PT treatment of cell proliferation disorders such as cancer  
 XX  
 PS Example 7; Page 149-151; 240pp; Japanese.  
 CC  
 CC This sequence represents a new protein containing PDZ domains encoded  
 CC by clone FH50, whose expression in human umbilical vascular endothelial  
 CC cells (HUVEC) is enhanced by stimulation with tumour necrosis factor  
 CC (TNF) alpha. The new protein is used to identify proteins which bind  
 CC to it (particularly to the PDZ domains) and the genes encoding them,  
 CC for use in the treatment of cell proliferation disorders such as cancer.  
 CC  
 XX  
 SQ Sequence 251 AA;  
 Query Match 22.2%; Score 1127; DB 20; Length 251;  
 Best Local Similarity 100.0%; Pred. No. 1.6e-86;  
 Matches 223; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 1 MLOWKSEFEERTINIAKNSSLGMYSAKNGIGMTVRSIIHGCAISRDRGAIAGCIL 60  
 |||||||  
 DB 29 mlgvskesferltlnakgnslgmtvsankdglmivrsllhgalsrdgialgdcil 88  
 OY 61 SINEESTISTVNAQARMLRRHSLIGPDIKITYVPARHLEEFKISLQOQSGRWALDIFS 120  
 |||||||  
 DB 89 slneestlsylnaagaramlrrhsligpdiklitypaeahleefkislqgsgvrmaldifs 148  
 OY 121 SYTGRIPELPEREGESESELONTAYSNMNPARYELMREPSKSLGISIVGSGWGSR 180  
 |||||||  
 DB 149 sytgripelperegeeseelqntaysnmqprvrlwrepskslqslivggrmggr 208  
 OY 181 LSNCEVNRGIFIKHVLDSPPAGKNGTLKPGDRIVEAPSOSESE 223  
 |||||||  
 DB 209 lsngevnrqgifikhvlcdspagkngtlkpgdriveapsgese 251  
 RESULT 11  
 AAY04739  
 ID AAY04739 standard; Protein: 272 AA.  
 AC  
 XX  
 AC AAY04739;  
 XX  
 DT 06-JUL-1999 (first entry)  
 XX  
 DE PDZ domain-containing protein gene encoded by clone FH850.  
 XX  
 KW PDZ domain; gene expression; human umbilical vascular endothelial cell;  
 KW HUVEC; stimulation; tumour necrosis factor; TNF; protein binding; PCR;  
 KW cell; proliferation disorder; cancer; primer; amplification.  
 XX  
 OS Homo sapiens.  
 OS  
 PN WO9907846-A1.  
 PN  
 PD 18-FEB-1999.  
 PD  
 XX 12-AUG-1998; 98WO-JP03603.  
 XX  
 PF 19-JUN-1998; 98JP-0189944.  
 PR 12-AUG-1997; 97JP-0230356.  
 PR  
 XX  
 PA (CHUG-) CHUGAI RES INST MOLECULAR MEDICINE INC.  
 PA  
 PI Funahashi S, Miyata S;  
 PI  
 XX WPI: 1999-167423/14.  
 DR

DR N-PSDB; AAX29970.  
 XX  
 PT Protein containing PDZ domain, whose expression is enhanced by TNF  
 PT stimulation - plays an important role in protein/protein  
 PT interactions and is used for screening for proteins for use in  
 PT treatment of cell proliferation disorders such as cancer  
 XX  
 PS Example 7; Page 152-154; 240pp; Japanese.  
 CC  
 CC This sequence represents a new protein containing PDZ domains encoded  
 CC by clone FH850, whose expression in human umbilical vascular endothelial  
 CC cells (HUVEC) is enhanced by stimulation with tumour necrosis factor  
 CC (TNF) alpha. The new protein is used to identify proteins which bind  
 CC to it (particularly to the PDZ domains) and the genes encoding them,  
 CC for use in the treatment of cell proliferation disorders such as cancer.  
 CC  
 XX  
 SQ Sequence 272 AA;  
 Query Match 21.4%; Score 1089; DB 20; Length 272;  
 Best Local Similarity 100.0%; Pred. No. 2.8e-83;  
 Matches 215; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 1 MLOWKSEFEERTINIAKNSSLGMYSAKNGIGMTVRSIIHGCAISRDRGAIAGCIL 60  
 |||||||  
 DB 29 mlgvskesferltlnakgnslgmtvsankdglmivrsllhgalsrdgialgdcil 88  
 OY 61 SINEESTISTVNAQARMLRRHSLIGPDIKITYVPARHLEEFKISLQOQSGRWALDIFS 120  
 |||||||  
 DB 89 slneestlsylnaagaramlrrhsligpdiklitypaeahleefkislqgsgvrmaldifs 148  
 OY 121 SYTGRIPELPEREGESESELONTAYSNMNPARYELMREPSKSLGISIVGSGWGSR 180  
 |||||||  
 DB 149 sytgripelperegeeseelqntaysnmqprvrlwrepskslqslivggrmggr 208  
 OY 181 LSNCEVNRGIFIKHVLDSPPAGKNGTLKPGDRIVE 215  
 |||||||  
 DB 209 lsngevnrqgifikhvlcdspagkngtlkpgdrive 243  
 RESULT 12  
 AAY04740  
 ID AAY04740 standard; Protein: 319 AA.  
 AC  
 XX  
 AC AAY04740;  
 XX  
 DT 06-JUL-1999 (first entry)  
 XX  
 DE PDZ domain-containing protein gene encoded by clone FH950.  
 XX  
 KW PDZ domain; gene expression; human umbilical vascular endothelial cell;  
 KW HUVEC; stimulation; tumour necrosis factor; TNF; protein binding; PCR;  
 KW cell; proliferation disorder; cancer; primer; amplification.  
 XX  
 OS Homo sapiens.  
 OS  
 PN WO9907846-A1.  
 PN  
 PD 18-FEB-1999.  
 PD  
 XX 12-AUG-1998; 98WO-JP03603.  
 XX  
 PF 19-JUN-1998; 98JP-0189944.  
 PR 12-AUG-1997; 97JP-0230356.  
 PR  
 XX  
 PA (CHUG-) CHUGAI RES INST MOLECULAR MEDICINE INC.  
 PA  
 PI Funahashi S, Miyata S;  
 PI  
 XX WPI: 1999-167423/14.  
 DR N-PSDB; AAX29971.  
 DR  
 PT Protein containing PDZ domain, whose expression is enhanced by TNF

stimulation - plays an important role in protein/protein interactions and is used for screening for proteins for use in treatment of cell proliferation disorders such as cancer

Example 7, Page 155-158; 240pp; Japanese.

This sequence represents a new protein containing PDZ domains encoded by clone FH950, whose expression in human umbilical vascular endothelial cells (HUVEC) is enhanced by stimulation with tumour necrosis factor (TNF) alpha. The new protein is used to identify proteins which bind to it (particularly to the PDZ domains) and the genes encoding them, for use in the treatment of cell proliferation disorders such as cancer.

Sequence 319 AA;

Query Match 21.4%; Score 1089; DB 20; Length 319;  
Best Local Similarity 100.0%; Pred. No. 3.7e-83;  
Matches 215; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLYNYSKSFERTINIANGNSSLGMTVSANKDGLMIVRSIIHGGAISRDRIAGDCTL 60  
DB 29 mlgvnskesfertlniakynsslgmtvsankdglmivrsllhggaistrdriagdcil 88  
QY 61 SINESTISVTNAQARAMLRRHSLIGPOIKITYVPAEHLKEFKISLGQSGRWALDIFS 120  
DB 89 sineestlsvtnaqaramlrrhsligpdkityvpaehleeekislsgsgqrvmaaldifs 148  
QY 121 SYTGNDIPELPEREGEGESELQNTAVSNMOPRRVELMREPSKSLGISIVGSGMGSR 180  
DB 149 sytgndipeelperegeegeseelqntaysnmqprvelwrepskslglsvgrvmgsr 208  
QY 181 LSNGEVMRGIFIKHVLDESPACKNGTLKPGDRIVE 215  
DB 209 lsngevrmrgiflkhvledspagknltlpgdrive 243

RESULT 13  
AAV04734

ID AAV04734 standard; Protein; 1239 AA.

AAV04734;

06-JUL-1999 (first entry)

Protein containing PDZ domain from clone 38-2-1c.

PDZ domain; gene expression; human umbilical vascular endothelial cell;  
HUVEC; stimulation; tumour necrosis factor; TNF; protein binding;  
cell; proliferation disorder; cancer.

Homo sapiens.

MO9907846-A1.

18-FEB-1999.

12-AUG-1998; 98WO-CP03603.

19-JUN-1998; 98JP-0189944.

12-AUG-1997; 97JP-0230356.

(CHUG- ) CHUGAI RES INST MOLECULAR MEDICINE INC.

Funahashi S, Miyata S;

WPI: 1999-167423/14.

N-PSDB; AAX29911.

Protein containing PDZ domain, whose expression is enhanced by TNF stimulation - plays an important role in protein/protein interactions and is used for screening for proteins for use in treatment of cell proliferation disorders such as cancer

Claim 1, Page 177-182; 240pp; Japanese.

This sequence represents a new protein containing a PDZ domain encoded by the clone 38-2-1c, whose expression in human umbilical vascular endothelial cells (HUVEC) is enhanced by stimulation with tumour necrosis factor (TNF) alpha. The new protein is used to identify proteins which bind to it (particularly to the PDZ domains) and the genes encoding them, for use in the treatment of cell proliferation disorders such as cancer.

Sequence 1239 AA;

Query Match 21.4%; Score 1089; DB 20; Length 1239;  
Best Local Similarity 100.0%; Pred. No. 3.5e-82;  
Matches 215; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLYNYSKSFERTINIANGNSSLGMTVSANKDGLMIVRSIIHGGAISRDRIAGDCTL 60  
DB 996 mlgvnskesfertlniakynsslgmtvsankdglmivrsllhggaistrdriagdcil 1055  
QY 61 SINESTISVTNAQARAMLRRHSLIGPOIKITYVPAEHLKEFKISLGQSGRWALDIFS 120  
DB 1056 sineestlsvtnaqaramlrrhsligpdkityvpaehleeekislsgsgqrvmaaldifs 1115  
QY 121 SYTGNDIPELPEREGEGESELQNTAVSNMOPRRVELMREPSKSLGISIVGSGMGSR 180  
DB 1116 sytgndipeelperegeegeseelqntaysnmqprvelwrepskslglsvgrvmgsr 1175  
QY 181 LSNGEVMRGIFIKHVLDESPACKNGTLKPGDRIVE 215  
DB 1176 lsngevrmrgiflkhvledspagknltlpgdrive 1210

RESULT 14  
AAV74151

ID AAV74151 standard; Protein; 206 AA.

AAV74151;

14-MAR-2000 (first entry)

Human prostate tumor EST fragment derived protein #338.

Pancreas; tumor; EST; expressed sequence tag; human; cytostatic;  
treatment.

Homo sapiens.

DE19820190-A1.

04-NOV-1999.

28-APR-1998; 98DE-1020190.

28-APR-1998; 98DE-1020190.

(META-) METAGEN GES GENOMFORSCHUNG MBH.

Rosenthal A, Specht T, Hinzmann B, Schmitt A, Pflersky C, Dahl E;

WPI: 1999-621386/54.

N-PSDB; AAZ52969.

New human nucleic acid sequences from pancreatic tumors, and related proteins -

Claim 23; Page 450; 502pp; German.

This invention describes novel polypeptides and their encoding nucleic acids derived from human pancreatic tumor tissue which have cytostatic activity. The sequences are also useful in producing pharmaceutical compositions for treatment of pancreatic tumors. AAV73814-VY4252





04-AUG-1998; 98US--0095321.  
PR 10-AUG-1998; 98US--0095325.  
PR 10-AUG-1998; 98US--0095916.  
PR 10-AUG-1998; 98US--0095929.  
PR 10-AUG-1998; 98US--0096012.  
PR 11-AUG-1998; 98US--0096143.  
PR 11-AUG-1998; 98US--0096146.  
PR 12-AUG-1998; 98US--0096329.  
PR 17-AUG-1998; 98US--0096757.  
PR 17-AUG-1998; 98US--0096766.  
PR 17-AUG-1998; 98US--0096768.  
PR 17-AUG-1998; 98US--0096773.  
PR 17-AUG-1998; 98US--0096791.  
PR 17-AUG-1998; 98US--0096867.  
PR 17-AUG-1998; 98US--0096891.  
PR 17-AUG-1998; 98US--0096894.  
PR 17-AUG-1998; 98US--0096895.  
PR 17-AUG-1998; 98US--0096897.  
PR 18-AUG-1998; 98US--0096949.  
PR 18-AUG-1998; 98US--0096950.  
PR 18-AUG-1998; 98US--0096959.  
PR 18-AUG-1998; 98US--0096960.  
PR 18-AUG-1998; 98US--0097022.  
PR 19-AUG-1998; 98US--0097141.  
PR 20-AUG-1998; 98US--0097218.  
PR 24-AUG-1998; 98US--0097661.  
PR 26-AUG-1998; 98US--0097951.  
PR 26-AUG-1998; 98US--0097952.  
PR 26-AUG-1998; 98US--0097954.  
PR 26-AUG-1998; 98US--0097955.  
PR 26-AUG-1998; 98US--0097971.  
PR 26-AUG-1998; 98US--0097974.  
PR 26-AUG-1998; 98US--0097978.  
PR 26-AUG-1998; 98US--0097979.  
PR 26-AUG-1998; 98US--0097986.  
PR 26-AUG-1998; 98US--0098014.  
PR 31-AUG-1998; 98US--0098525.  
PR 16-SEP-1998; 98US--0100634.  
PR 12-JAN-1999; 99US--0115565.

(GENE ) GENENTECH INC.

Baker K, Chen J, Goddard A, Gurney AL, Smith V, Watanabe CK,  
PI Wood WT, Yuan J;  
XX  
XX  
PT N-PSDB; AAZ65028.  
DR  
XX  
XX  
XX

Membrane-bound proteins and related nucleotide sequences -  
  
claim 12; Fig 147; 822pp; English.

The invention provides membrane-bound PRO polypeptides and  
CC polynucleotides encoding them. The PRO sequences of the invention were  
CC identified based on extracellular domain homology screening. The PRO  
CC sequences have homology with proteins including LDL receptors, TIE  
CC ligands and various enzymes. The membrane-bound proteins and receptor  
CC molecules are useful as pharmaceutical and diagnostic agents. Receptor  
CC immunoadhesins, for instance, can be used as therapeutic agents to block  
CC receptor-ligand interactions. The membrane-bound proteins can also be  
CC employed for screening of potential peptide or small molecule inhibitors  
CC of the relevant receptor/ligand interaction. The PRO encoding sequences  
CC are useful as hybridization probes, in chromosome and gene mapping and in  
CC the generation of antisense RNA and DNA. PRO nucleic acid sequences  
CC will also be useful for the preparation of PRO polypeptides, especially  
CC by recombinant techniques.

Sequence 632 AA:  
XX

|                       |                   |                    |             |             |
|-----------------------|-------------------|--------------------|-------------|-------------|
| Query Match           | 11.58;            | Score 583;         | DB 21;      | length 632; |
| Best Local Similarity | 25.68;            | Pred. No. 4.4e-40; |             |             |
| Matches 195;          | Conservative 119; | Mismatches 254;    | Indels 194; | Gaps 21.    |

|    |     |  |      |
|----|-----|--|------|
| 0Y | 281 | GELHMI--ELTEKGHSGLSLAGKNGSRMSVFIWIDPNGAAGKORQJADELLEING      | 338  |
| Db | 22  | gnhlflyselckgashgylt---kdkfrts-----qdg-----                  | 51   |
| 0Y | 339 | QILYGRSHONASSIICKAPSKVKIIFIRNKDAVQMAVCBNAVEPRLPSPSENENKETE   | 398  |
| Db | 52  | -----cpdgasatlataps-----pevasaatslmtde                       | 80   |
| 0Y | 399 | PYVTITSDAAYDLSEFKNVQHLLEPKDOGGLCIAISEDLSGVILKSLTEHGAANDGRL   | 458  |
| Db | 81  | pgl---dnpayvsaadqbpalspvdsgsrntrrparfetslrtsfkk-----         | 127  |
| 0Y | 459 | KVGGQILAAVDEIYGVPIEKFILSLTAKNTVTLTHAENPDQAVPSAAGAASGKN       | 518  |
| Db | 128 | -----inatsvlrttcksgsavanhd-----                              | 149  |
| 0Y | 519 | SSQSLWVQSGSPPEESIRNTRSSTPAIFASDAPCPIIPGCE-TWIEISK--GRTLG     | 575  |
| Db | 150 | -----qgremsentapevf---prlyhllpdsgetstsklnrvpsasls            | 191  |
| 0Y | 576 | LSIVGSDTLIGAFITHEVEEAGACKDGRIMACDQILEVNGIDLKRAITHAENIVLRQT   | 635  |
| Db | 192 | lrlvgsetpvlhllqhllydgvlaydgrlllpdglilkvngmdlsmvphayvrltrp    | 251  |
| 0Y | 636 | PQRVRLTIRD-----EAP--YKEEENCOPLTIELQ-KRGKGLGSLIVGRKNDTG       | 688  |
| Db | 252 | qcvlwlvtmteqkfrtrsmngqaradaytrd--dsfhyvlmspspeeqgylkvlrvkep  | 309  |
| 0Y | 684 | VFVSDIVYGFIADPDRGLRQDQDQILVNGEDBVRNAISOEAVALLKCSIGVTTLEVRKA  | 743  |
| Db | 310 | vflfnvldgvyahngqleendrclalnhpdlrygspeasaahlqaseerrvhlvavr--- | 366  |
| 0Y | 744 | GPFHSERRPSOTSOVSEBSLSFTFPLPSGSSSTSSLESSSKKNLASEIQ-GLATVEKK   | 807  |
| Db | 367 | -----qvrtrspdlfgeaawmsngswspargersnprkplhlttchevvnalqk       | 416  |
| 0Y | 803 | GPTDSLGISIGVGGSPRLGDVPIFLAMNHGVAQAQOKLWVGCRITYICTSTREGMHT    | 865  |
| Db | 417 | dpgeelgtmctvagasahrewdrlprlylvevrgvlsrgrlktgdlilnvdvalewvrs  | 477  |
| 0Y | 863 | QAVNLKNAAGSIEMOVAVGADVVTGHHQBPASSLSFTGLTSTSIPODGLAP---       | 918  |
| Db | 477 | eavalllkrtssslvka-----levkeyqedeccspalds-----nhmmaprpdws     | 524  |
| 0Y | 919 | -----QCKSTLERPGD--LGSIVYGGSFPHGDLPIYKVTYFAKGAASED            | 965  |
| Db | 525 | pswvwmjelprclynckdlvtrntagslglcivgyveeyngnkrpfikslvleglpaynd | 584  |
| 0Y | 964 | GLRKGDOITIAVNGOSLEGVTHHEAVNALIKRTKGTVMUWS                    | 1005 |
| Db | 585 | grlrcgdlllavngtrstsgmhacrlarllkelygrlctltvs                  | 626  |

Search completed: July 12, 2001, 14:40:37  
Job time: 200 sec

Fri Jul 13 15:00:10 2001

us-09-502-698-2.rag

Page 14

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw-model

Run on: July 12, 2001, 14:41:51 ; Search time 40.62 Seconds  
(without alignments)  
1884.673 Million cell updates/sec

Title: US-09-502-698-2

Perfect score: 5085

Sequence: 1 MGNVSKESFEFRTINAKGN.....EEAVALKRKGVTLVWLS 1005

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score  | Query Match | Length | DB ID    | Description          |
|------------|--------|-------------|--------|----------|----------------------|
| 1          | 4413.5 | 86.8        | 2054   | 2 T46612 | multi PDZ domain p   |
| 2          | 4314.5 | 84.8        | 2055   | 2 T30259 | multiple PDZ domai   |
| 3          | 814    | 16.0        | 2172   | 2 T20145 | hypothetical prote   |
| 4          | 589    | 11.6        | 728    | 2 T09457 | numb-binding prote   |
| 5          | 586    | 11.5        | 628    | 2 T09458 | numb-binding prote   |
| 6          | 534    | 10.5        | 1012   | 2 T23160 | hypothetical prote   |
| 7          | 523    | 10.0        | 2450   | 2 S71625 | protein-tyrosine-p   |
| 8          | 509.5  | 10.0        | 2466   | 2 T67629 | protein-tyrosine p   |
| 9          | 501.5  | 9.9         | 2294   | 2 T54971 | protein tyrosine p   |
| 10         | 494.5  | 8.9         | 1256   | 2 T08209 | protein-tyrosine-p   |
| 11         | 453.5  | 8.0         | 852    | 2 T10811 | brain-specific ang   |
| 12         | 407.5  | 8.0         | 852    | 2 T10811 | channel associated   |
| 13         | 399.5  | 7.9         | 1131   | 2 T15617 | hypothetical prote   |
| 14         | 397.5  | 7.8         | 870    | 2 G01974 | channel associated   |
| 15         | 395    | 7.8         | 1277   | 2 T14152 | synaptic scaffold    |
| 16         | 385    | 7.6         | 1171   | 2 T42372 | probable guanylate   |
| 17         | 384    | 7.6         | 767    | 2 T09559 | postsynaptic densi   |
| 18         | 383    | 7.5         | 724    | 2 JH0800 | postsynaptic densi   |
| 19         | 380    | 7.5         | 720    | 2 A45436 | synapse-associated   |
| 20         | 379    | 7.5         | 960    | 2 A39651 | discs-large tumor    |
| 21         | 378.5  | 7.4         | 911    | 2 T56552 | synapse-associated   |
| 22         | 377    | 7.4         | 904    | 2 T18757 | homolog of Drosoph   |
| 23         | 377    | 7.4         | 926    | 2 T32733 | homolog of Drosoph   |
| 24         | 357    | 7.0         | 1112   | 2 T32733 | AMPA glutamate rec   |
| 25         | 352    | 6.9         | 1337   | 2 T13948 | atypical protein k   |
| 26         | 348    | 6.8         | 1464   | 2 T13716 | cathepsin gene prote |
| 27         | 264    | 5.2         | 1360   | 2 T34302 | cell polarity prot   |
| 28         | 246    | 4.8         | 1163   | 2 JEO366 | tight junction pro   |
| 29         | 243    | 4.8         | 1736   | 2 A47747 | tight junction pro   |

|    |       |     |      |          |                    |
|----|-------|-----|------|----------|--------------------|
| 30 | 237.5 | 4.7 | 1745 | 2 A46431 | tight junction-ass |
| 31 | 228   | 4.5 | 1116 | 2 T54378 | gene X104 protein  |
| 32 | 216   | 4.2 | 1367 | 2 S60315 | PSD-95-related pro |
| 33 | 210   | 4.1 | 1367 | 2 T13703 | tama protein - fru |
| 34 | 197   | 3.9 | 505  | 2 S62894 | alpha-synuclein -  |
| 35 | 194.5 | 3.8 | 126  | 2 T81210 | tyrosine phosphata |
| 36 | 190.5 | 3.7 | 1829 | 2 T41751 | 1-afadin - rat     |
| 37 | 188   | 3.7 | 1281 | 2 T00346 | hypothetical prote |
| 38 | 187   | 3.7 | 723  | 2 T14765 | hypothetical prote |
| 39 | 187   | 3.7 | 1095 | 2 T43275 | neurabin - rat     |
| 40 | 187   | 3.7 | 1663 | 2 T42092 | s-afadin - rat     |
| 41 | 187   | 3.7 | 1893 | 2 A56158 | eye development pr |
| 42 | 186.5 | 3.7 | 1117 | 2 T81209 | tyrosine phosphata |
| 43 | 186.5 | 3.7 | 1666 | 2 T43169 | hypothetical prote |
| 44 | 186.5 | 3.7 | 5627 | 2 C83359 | hypothetical prote |
| 45 | 185.5 | 3.6 | 488  | 2 T51379 | synuclein - Pacif  |

## ALIGNMENTS

## RESULT 1

T46612 multi PDZ domain protein 1 - rat

C:Species: Rattus norvegicus (Norway rat)

C:Date: 18-Feb-2000 #sequence\_revision 18-Feb-2000 #text\_change 21-Jul-2000

R:Accession: T46612

A:Title: Cloning and characterization of MUPP1, a novel PDZ domain protein.

A:Reference number: 223104; MUID:98196865

A:Accession: T46612

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-2054 <URL>

A:Cross-references: EMBL:AJ001320; NID:g2959978; PIDN:CAA04681.1; PID:g2959979

A:Experimental source: brain

C:Genetics:

A:Gene: MUPP1

|                       |       |   |                                  |                                   |
|-----------------------|-------|---|----------------------------------|-----------------------------------|
| Query Match           | 86.8% | Score 4413.5;   | DB 2;                            | Length 2054;                      |
| Best local similarity | 81.5% | Pred. No. 5.1e-244;                                       |                                  |                                   |
| Matches               | 875;  | Conservative  | 64;                              | Mismatches 62; Indels 73; Gaps 2; |
| QY                    | 2     | LGNVSKESFEFRTINAKGN                                       | SLGNTVSAKNDGLMIVRSIIHGAI         | SRDRIAGDCILS 61                   |
| DB                    | 984   | LQSMQEAFFERTVTAKSSSLGNTVSAKNDGLMIVRSIIHGAI                | SRDRIAGDCILS                     | 1043                              |
| QY                    | 62    | INEESTISVTNQAARMLRRHS                                     | SLGPDKITVPAEHLEEFKISGQSGRMALDISS | 1103                              |
| DB                    | 1044  | INEESTISVTNQAARMLRRHS                                     | SLGPDKITVPAEHLEEFKISGQSGRMALDISS | 1103                              |
| QY                    | 122   | YTGRIPELPEREGEGESELQNTAYSNMNQPRRVELMRREPSKSLGSI           | YVGRMGSRNL 181                   |                                   |
| DB                    | 1104  | YTGRIPELPEREGEGESELQNTAYSNMNQPRRVELMRREPSKSLGSI           | YVGRMGSRNL 1163                  |                                   |
| QY                    | 162   | SNGEVVRGIFIKHVLIEDSPAKNGTLKPDRIE                          | -----                            | 215                               |
| DB                    | 1164  | SNGEVVRGIFIKHVLIEDSPAKNGTLKPDRIE                          | -----                            | 215                               |
| QY                    | 216   | -----   | -----                            | -----                             |
| DB                    | 1224  | FWQSVIVNRPRKSPPLSLPHSLYPKCSSTTPFAESLQITSDKAPSSQSESEKATICS | 1283                             |                                   |
| QY                    | 232   | VPPPPSAFAEMGSDHTOSASAKISQDYDKDEDFGYSKNI                   | RERYGTLTGLHMLEKSG 291            |                                   |
| DB                    | 1284  | VPPSSPVSEKSSDYAQPSATVADEDDKDEDFGYSKNI                     | QERYGTLTGLHMLEKSG 1343           |                                   |
| QY                    | 232   | HSGGLSLAGNKDRSRMSVFIYIDPNGAAGKDLQIADLELE                  | INGOILYGRSQNASS 351              |                                   |
| DB                    | 1344  | HSGGLSLAGNKDRSRMSVFIYIDPNGAAGKDLQIADLELE                  | INGOILYGRSQNASS 1403             |                                   |



C:Species: *Caenorhabditis elegans*  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 29-Oct-1999  
C:Accession: T20145; T25009  
R:Stulston, J.  
submitted to the EMBL Data Library, November 1994  
A:Reference number: Z19229  
A:Accession: T20145  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1\*2172 <MIL>  
A:Cross-references: EMBL:Z46792; PIDN:CAA86769.1; GSPDB:GN00020; CESP:C52A11.4  
A:Experimental source: clone C52A11  
R:Coles, L.  
submitted to the EMBL Data Library, November 1994  
A:Reference number: Z19968  
A:Accession: T25009  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-2172 <WT>  
A:Cross-references: EMBL:Z46795; PIDN:CAA86789.1; GSPDB:GN00020; CESP:C52A11.4  
A:Experimental source: clone T19E10  
C:Genetics:  
A:Gene: CESP:C52A11.4  
A:Map position: 2  
A:Introns: 2/2; 78/1; 111/3; 156/3; 208/1; 230/3; 290/2; 341/1; 389/3; 431/1; 461/3; 632/3; 1747/1; 1946/3; 2049/3; 2096/3; 2145/3

```

Query Match Similarity 16.0%: Score 814; DB 2; Length 2172;
Match Local Similarity 20.9%: Pred. No. 4e-38;
Matches 307; Conservative 137; Mismatches 239; Indels 768; Gaps 33;

QY 11 ERTIANGNSSIGMTVSNKD-GI-GMIVRSIIIGCAISRDRGAIADGILSINEESTI 68
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 859 ERTVKGALPGLAVLGDGDKDGVNCGVKSICGKAAVADGRIVQVDFTKINTESLR 918
QY 69 SYTNQAARMLRHRSLIGPIDIKITYVP----- 95
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 919 NVTNSQARAILKRLNVLGFCNVYITISADAKTWKRFQRPDESSPIINRLSPKVPKF 978
QY 96 -----AEHLFEFKISLG-----QOSGRVM 114
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 979 YRSPMQOESQSKTEMTDDTEAPRSIMTDSMSEHIIKTFDLAEGSSRSHHDEQVRNMS 1038
QY 115 AL-----DIF 119
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 1039 RLIDVEVDDEVNLIKEATIDATIELRVLKTKDKMSNCNKRERLESPPLPPPEVL 1098
QY 120 SS-----YTGRI-----PELPE 132
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 1099 SSPKSPVASVQPTPRDLEEVLTBSTASSLEHSGQRTSQLHILSTBEEVLAQATPPSSPE 1158
QY 133 REEGE----- 137
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 1159 NKSEVPEPSISPGIKLAGEVTAPEIEYVKAQENVDRAETAATGAEEATSTPTAEAI 1218
QY 138 -----GESEELONTA-----Y 148
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 1219 GNSKODESTTISISQOSVGLQTOALNSTEVENMSRVTSRPSTGES-LONQAQOLVRS 1277
QY 149 SNMNOPRRVELMRPEPSKSLGISIVGCR---GMSGRLNGEVMRGIFIKHVLDESDPAGN 204
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 1278 KVMGEARFVTLVLRERNKSPGSGISIVGGRVEVSGKGLPQTGNVYCGIIFKSVLPNPAARS 1337
QY 205 GTLKPGDRIVEAP----- 217
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 1338 GOMNNGDRIVSYNDVDLRDATHQAVNMAIKNASNPREVLOSLHTNOONMINSASNSTVG 1397
QY 218 -----SSESEPEKAPLCSVPP-----PPPS---AFAEMGSDHTOS 250
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 1398 SVRFENAKREDEELPTTALVTPFLKPMISSSGSSTSKPANNFPPPSISTTTTTSMESEKDE 1457
QY 251 SASKISODVDED----- 263

```

**RESULTS**

**A:** Molecule type: mRNA  
Accession: T09457  
Title: The mammalian Numb phosphotyrosine binding domain: Characterization of blind

**B:** Reference number: 216678; MUID:98204916

**C:** Species: Mus musculus (house mouse)  
Citation: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 02-Sep-2000  
Accession: T09457

**D:** R.Dho, S.E.; Jacob, S.; Wolting, C.D.; French, M.B.; Rohrschneider, L.R.; McGlade, C.  
J. Biol. Chem. 273, 9179-9187, 1998

**E:** Accession: T09457

**F:** Status: preliminary; translated from GB/EMBL/DDBJ

**G:** A:Molecule type: mRNA

**H:** Accession: T09457

**I:** Title: The mammalian Numb phosphotyrosine binding domain: Characterization of blind

**J:** Reference number: 216678; MUID:98204916

**K:** Species: Mus musculus (house mouse)

**L:** Citation: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 02-Sep-2000

**M:** Accession: T09457

**N:** R.Dho, S.E.; Jacob, S.; Wolting, C.D.; French, M.B.; Rohrschneider, L.R.; McGlade, C.

**O:** J. Biol. Chem. 273, 9179-9187, 1998

**P:** Accession: T09457

**Q:** Status: preliminary; translated from GB/EMBL/DDBJ

**R:** A:Molecule type: mRNA

**S:** Accession: T09457

**T:** Title: The mammalian Numb phosphotyrosine binding domain: Characterization of blind

**U:** Reference number: 216678; MUID:98204916

**V:** Species: Mus musculus (house mouse)

**W:** Citation: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 02-Sep-2000

**X:** Accession: T09457

**Y:** R.Dho, S.E.; Jacob, S.; Wolting, C.D.; French, M.B.; Rohrschneider, L.R.; McGlade, C.

**Z:** J. Biol. Chem. 273, 9179-9187, 1998

**AA:** Accession: T09457

**AB:** Status: preliminary; translated from GB/EMBL/DDBJ

**AC:** A:Molecule type: mRNA

**AD:** Accession: T09457

**AE:** Title: The mammalian Numb phosphotyrosine binding domain: Characterization of blind

**AF:** Reference number: 216678; MUID:98204916

**AG:** Species: Mus musculus (house mouse)

**AH:** Citation: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 02-Sep-2000

**AI:** Accession: T09457

**AJ:** R.Dho, S.E.; Jacob, S.; Wolting, C.D.; French, M.B.; Rohrschneider, L.R.; McGlade, C.

**AK:** J. Biol. Chem. 273, 9179-9187, 1998

**AL:** Accession: T09457

**AM:** Status: preliminary; translated from GB/EMBL/DDBJ

**AN:** A:Molecule type: mRNA

**AO:** Accession: T09457

**AP:** Title: The mammalian Numb phosphotyrosine binding domain: Characterization of blind

**AQ:** Reference number: 216678; MUID:98204916

**AR:** Species: Mus musculus (house mouse)

**AS:** Citation: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 02-Sep-2000

**AT:** Accession: T09457

**AU:** R.Dho, S.E.; Jacob, S.; Wolting, C.D.; French, M.B.; Rohrschneider, L.R.; McGlade, C.

**AV:** J. Biol. Chem. 273, 9179-9187, 1998

**AW:** Accession: T09457

**AX:** Status: preliminary; translated from GB/EMBL/DDBJ

**AY:** A:Molecule type: mRNA

**AZ:** Accession: T09457

**BA:** Title: The mammalian Numb phosphotyrosine binding domain: Characterization of blind

**BB:** Reference number: 216678; MUID:98204916

**BC:** Species: Mus musculus (house mouse)

**BD:** Citation: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 02-Sep-2000

**BE:** Accession: T09457

**BF:** R.Dho, S.E.; Jacob, S.; Wolting, C.D.; French, M.B.; Rohrschneider, L.R.; McGlade, C.

**BG:** J. Biol. Chem. 273, 9179-9187, 1998

**BH:** Accession: T09457

**BI:** Status: preliminary; translated from GB/EMBL/DDBJ

**BJ:** A:Molecule type: mRNA

**BK:** Accession: T09457

**BL:** Title: The mammalian Numb phosphotyrosine binding domain: Characterization of blind

**BM:** Reference number: 216678; MUID:98204916

**BN:** Species: Mus musculus (house mouse)

**BO:** Citation: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 02-Sep-2000

**BP:** Accession: T09457

**BQ:** R.Dho, S.E.; Jacob, S.; Wolting, C.D.; French, M.B.; Rohrschneider, L.R.; McGlade, C.

**BR:** J. Biol. Chem. 273, 9179-9187, 1998

**BS:** Accession: T09457

**BT:** Status: preliminary; translated from GB/EMBL/DDBJ

**BU:** A:Molecule type: mRNA

**BV:** Accession: T09457

**BW:** Title: The mammalian Numb phosphotyrosine binding domain: Characterization of blind

**BX:** Reference number: 216678; MUID:98204916

**BY:** Species: Mus musculus (house mouse)

**BZ:** Citation: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 02-Sep-2000

**CA:** Accession: T09457

**CB:** R.Dho, S.E.; Jacob, S.; Wolting, C.D.; French, M.B.; Rohrschneider, L.R.; McGlade, C.

**CC:** J. Biol. Chem. 273, 9179-9187, 1998

**CD:** Accession: T09457

**CE:** Status: preliminary; translated from GB/EMBL/DDBJ

**CF:** A:Molecule type: mRNA

**CG:** Accession: T09457

**CH:** Title: The mammalian Numb phosphotyrosine binding domain: Characterization of blind

**CI:** Reference number: 216678; MUID:98204916

**CJ:** Species: Mus musculus (house mouse)

**CK:** Citation: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 02-Sep-2000

**CL:** Accession: T09457

**CM:** R.Dho, S.E.; Jacob, S.; Wolting, C.D.; French, M.B.; Rohrschneider, L.R.; McGlade, C.

**CN:** J. Biol. Chem. 273, 9179-9187, 1998

**CO:** Accession: T09457

**CP:** Status: preliminary; translated from GB/EMBL/DDBJ

**CQ:** A:Molecule type: mRNA

**CR:** Accession: T09457

**CS:** Title: The mammalian Numb phosphotyrosine binding domain: Characterization of blind

**CT:** Reference number: 216678; MUID:98204916

**CU:** Species: Mus musculus (house mouse)

**CV:** Citation: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 02-Sep-2000

**CV:** Accession: T09457

**CV:** R.Dho, S.E.; Jacob, S.; Wolting, C.D.; French, M.B.; Rohrschneider, L.R.; McGlade, C.

**CV:** J. Biol. Chem. 273, 9179-9187, 1998

**CV:** Accession: T09457

**CV:** Status: preliminary; translated from GB/EMBL/DDBJ

**CV:** A:Molecule type: mRNA

**CV:** Accession: T09457

**CV:** Title: The mammalian Numb phosphotyrosine binding domain: Characterization of blind

**CV:** Reference number: 216678; MUID:98204916

**CV:** Species: Mus musculus (house mouse)

**CV:** Citation: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 02-Sep-2000

**CV:** Accession: T09457

**CV:** R.Dho, S.E.; Jacob, S.; Wolting, C.D.; French, M.B.; Rohrschneider, L.R.; McGlade, C.

**CV:** J. Biol. Chem. 273, 9179-9187, 1998

**CV:** Accession: T09457

**CV:** Status: preliminary; translated from GB/EMBL/DDBJ

**CV:** A:Molecule type: mRNA

**CV:** Accession: T09457

**CV:** Title: The mammalian Numb phosphotyrosine binding domain: Characterization of blind

**CV:** Reference number: 216678; MUID:98204916

**CV:** Species: Mus musculus (house mouse)

**CV:** Citation: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 02-Sep-2000

**CV:** Accession: T09457

**CV:** R.Dho, S.E.; Jacob, S.; Wolting, C.D.; French, M.B.; Rohrschneider, L.R.; McGlade, C.

**CV:** J. Biol. Chem. 273, 9179-9187, 1998

**CV:** Accession: T09457

**CV:** Status: preliminary; translated from GB/EMBL/DDBJ

**CV:** A:Molecule type: mRNA

**CV:** Accession: T09457

**CV:** Title: The mammalian Numb phosphotyrosine binding domain: Characterization of blind

**CV:** Reference number: 216678; MUID:98204916

**CV:** Species: Mus musculus (house mouse)

**CV:** Citation: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 02-Sep-2000

**CV:** Accession: T09457

**CV:** R.Dho, S.E.; Jacob, S.; Wolting, C.D.; French, M.B.; Rohrschneider, L.R.; McGlade, C.

**CV:** J. Biol. Chem. 273, 9179-9187, 1998

**CV:** Accession: T09457

**CV:** Status: preliminary; translated from GB/EMBL/DDBJ

**CV:** A:Molecule type: mRNA

**CV:** Accession: T09457

**CV:** Title: The mammalian Numb phosphotyrosine binding domain: Characterization of blind

**CV:** Reference number: 216678; MUID:98204916

**CV:** Species

A:Residues: 1-728 <DHO>  
A:Cross-references: EMBL:AF034745; NID:g3041878; PIRN:AAC40075.1; PIRID:g3041879  
A:Experimental source: developmental stage: embryo  
C:Genetics:  
A:Gene: lux  
C:Superfamily: RING\_finger\_homology  
F:41-88/domain: RING\_finger\_homology <RRN>

|                       |                   |                   |             |             |
|-----------------------|-------------------|-------------------|-------------|-------------|
| Query Match           | 11.6%             | Score 589;        | DB 2;       | Length 728; |
| Best Local Similarity | 29.1%             | Pred No. 5.4e-26; |             |             |
| Matches 188;          | Conservative 115; | Mismatches 238;   | Indels 106; | Gaps 23;    |

|    |     |   |      |
|----|-----|---|------|
| QY | 411 | SSFKNVHLELPKD-----QGLG--GIAISEETDGVILKSTLEHGAVTADRLKVGQD      | 463  |
| Db | 130 | TJSCAKASHGLKDKRRKRGDQDGCPOGCASLMA7TLLS-----EYVSAANT-----      | 173  |
| QY | 464 | LLAVDEIVGVPIPEKIFSLKLTAKTIVKLTJHAENPDSOANP---SAGCAASEKKNS     | 520  |
| Db | 174 | ISLMDDEGLDNP--AYVSVDEGEVPVANSSDGSRNRTARPERFESTYMSRFFKINRA     | 231  |
| QY | 521 | QS-LWVPOSGSPESPI-----RNTSRKSTPAIFASPPACPIPGCE--TTEISKGTGL     | 574  |
| Db | 232 | LSALKRTSGSVYVANHVDGGRDNSENTEYEVF---PRLFHLIPDEIISIKINRADPSE    | 268  |
| QY | 575 | GLST--VGGSDPLLCAFTTIEHYEEGACAKOGRLNAGDQDLEVNGIDLKRTATDEALNVL  | 632  |
| Db | 289 | SLSTLPLVGSEPEPLNIIIIQHTLRDGVILARDRLLEGDIIILKNGANDISVPHNYVRL   | 348  |
| QY | 633 | ROPQVRVLTLYRDE-----APYKEEYCDTTLTIELOK-KPKGGLSLIVGKRNDT        | 682  |
| Db | 349 | ROPQVRLTLYREQKFRSRNANHPDYSYGRPDSDPHVILINKSPREOGLIKLVRVDEP     | 408  |
| QY | 683 | GVFVSDIYKGGIADBDGRITDGDQDITLVNGEDVYRNASOEAVALLCSLTGTVLVEGR    | 742  |
| Db | 409 | GVFTINVLNGVADNRKGLLEENDRYALINGHDIRFESPEAHLLDASERHNLVYSR--     | 466  |
| QY | 743 | AGPFHSERRPS-----QTSQVSEGLSLSTFFPLSGSTSESLSSKKNALASEIOGLRT     | 797  |
| Db | 467 | -----QVROSSPDIEFOEAGWISNGOOS-----PGRGER-----NTASKPATACHE---KV | 508  |
| QY | 798 | VEMKGPRTDLSLISLAGVGSPLGDPVPIFIAMNHPYGVAAOQKLVGDRVITVIGNSTE    | 857  |
| Db | 509 | VSVKMDPSESJLGMVYGGGASHREMDLPIVIVISYERPGVLSRDRKITGDIILVANGIELT | 568  |
| QY | 858 | GMTHTQAVNLKNAGSISIEVYVAGDVSVTGHNDEPASSLSLSTGTLSTSTIFODDLP     | 917  |
| Db | 569 | EVSRTAEVAAILKSAPSSVVLKAL---EYK-----EOEAQEDCSPALDLS---NNHVT    | 615  |
| QY | 918 | P-----OQKSTLERGRPDG-LGFSIVYGGVGSPPGADPIYVYKTFAGK              | 958  |
| Db | 616 | PGDMSPSVMWMLLEPOYLCLNCCKDVLIRNTAGSLGCGIVGGEYEGSNKRPFLKSTIVEGT | 675  |
| QY | 959 | AAASEDGLKRGDQITAVNGQSLSEGVTHIEAVAILKRTKSTVTLMLVLS             | 1005 |
| Db | 676 | PAYNNGRIRCGDILLVNGRSTSGMIHACLARMLKELKRIITLLTAS                | 722  |

```

RESULT      5
T09458
numb-binding protein LNXp70 - mouse
C:Species: Mus musculus (house mouse)
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jul-2000
C:Accession: T09458
R:Dho, S.E.; Jacob, S.; Woltung, C.D.; French, M.B.; Rohrschneider, L.R.; McGlade, C.J.
J. Biol. Chem. 273, 9179-9187, 1998
A:Title: The mammalian Numb phosphotyrosine binding domain: Characterization of binding
A:Reference number: z16678; MUID:98204916
A:Accession: T09458
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-628 <DHO>
A:Cross-references: EMBL:AF034746; NID:g3041880; PIDN:AAC40076.1; PID:g3041881

```

A;Experimental source: brain  
C;Genetics:  
A;Gene: 1nx

|                       |                   |                    |             |             |
|-----------------------|-------------------|--------------------|-------------|-------------|
| Query Match           | 11.5%;            | Score 586;         | DB 2;       | Length 628; |
| Best Local Similarity | 29.0%;            | Pred. No. 6.4e-26; |             |             |
| Matches 187;          | Conservative 114; | Mismatches 237;    | Indels 106; | Gaps 23;    |

```

OY 414 KNOVHLELPKD-----QOGL-GIASEEDTSGYIKSLFPHGVAADGRKLKVDQJLA 466
Db 33 KGASHYGLTKORRKSQOCPCPDGASLMATLIS-----PEVSAAI-----ISL 76
OY 467 VDDELIVGYPLEKFIISLTKAKMTVKLTITHAENDSQAVP---SAAGAAGEKKNSSQS- 522
Db 77 MTDEGCLNPP--AYVSSYDEDEPVANSSDGSKNTRRAPFERSTMRSRSEFKIKRALSA 134
OY 523 LAMVQSGSPPEEST-----RNTSRSSPTALFASDPATCIIPGCE-TTIEISKGRIGLGS 577
Db 135 LRRKTSGGSVANHNDOGRDNSEMTVPPEVF---PLRFLHPIDGELTSIKINNADSELS 191
OY 578 I--VGSGPTLLGAFTIHEVEYEGACCKGRMLWAGQOILEVNGIDIRKATTHDEAINVLRQT 635
Db 192 IRLVGSGTPLVHTIIQIHTYRDGVYARQGRLLPGDIIILKVNMDISNPNHNAVALLQOP 251
OY 636 PORVRLTYRDE-----APYKEEVCDTLTIELOK-KPGKGLGSLIVKRNQTVF 685
Db 252 CQVRLATFLRQCKFRSRGNAHVPDSYGPBDDSFHYILKNKSSPEEDGLTKLVRVDEPEVF 311
OY 686 VSDIYKGGIADPDGLIGDQDQILVNGSDVNRASQEAVALIKGSLGTVTEVGRKAGP 745
Db 312 IFNVNLGVADRBQGLENDKRYLAINGHDLRPGSPESAAHLIQASERYVHLVSR----- 366
OY 746 FHSERRPS-----QTSQVSEGLSSFTFPLGSGSTSESLESSSKKNALASEIQLRTYEM 800
Db 367 --QVQSSPDIFQEGAGWISNGQOS---PGGGER-----NTASKRAAATCHE---KVYSV 411
OY 801 KKGPDLSGISIAGVGSPGLDVPFIAMMPTGVAQATOKRLVYDRTVYTIGSTBEAMT 860
Db 412 WKDSESGMTVVGGAHSREMDPLTYIVLSVEGGYISDGRKTKTDILLNNGLELTVES 471
OY 861 HTOAVNLKKNASGSIEMQVAVAGDVSVYTGHHQEPASSLSFTGLTSTISIFODDGLSP-- 918
Db 472 RTEAVAILKSAPESSVYKAL---EVK-----EQDAQEDCSPAALDS---NHNTTPRGD 518
OY 919 -----OCKSITLERGPRG-LGFSIVGSGSPHGDLPLYVTYVFAKGAAS 961
Db 519 WSPSNVMMLEPOYLCKNKDVILRKNTGSLGFCFVCGVEEYSGNKPFFIKSIVGCFPAY 578
OY 962 EDGRUKRGDQILIAVNGOSLEGVTHEEAVAILKRTGVTLWMLS 1005
Db 579 NDGRIRCGDILLAVNGKSTGMIHCLARMLKELKGRITLLTAS 622

```

RESULT 6  
T23160  
hypochemical protein K01A6.1 - *Caenorhabditis elegans*  
C:Species: *Caenorhabditis elegans*  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 18-Aug-2000  
C:Accession: T23160  
R:Cottage, A.  
submitted to the EMBL Data Library, January 1996  
A:Reference number: Z19701  
A:Accession: T23160  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Kolecule type: DNA  
A:Residues: 1..1012 <MIL>  
A:Cross-references: EMBL:Z68750; PIDN:CAA92965.1; GSPDB:GN00022; CESP:K01A6.1  
A:Experimental source: clone K01A6  
C:Genetics:  
A:Gene: CESP:K01A6.1  
A:Map position: 4  
A:Introns: 107/1; 151/2; 183/2; 270/2; 316/3; 410/3; 628/3; 922/3; 971/3

C:Superfamily: WW repeat homology  
F:131-168/Domain: WW repeat homology <WWR>

Query Match 10.5%; Score 534; DB 2; Length 1012;  
Best Local Similarity 24.1%; Pred. No. 1.3e-22;  
Matches 208; Conservative 123; Mismatches 298; Indels 224; Gaps 31;

```

OY 204 NGLTKRGRIVAPPSQSESEPEKAPLCSP-----PPPSAFAMGSDHTQSSAKS 256
DB 317 SCLTK-----SSSSPPDSFSSPTRYRKFGDPPEATTSADYDHSMKFSRS 365
OY 257 QVDEDEDEGYSMKNIREYGTGLHMELEKSGSLGSLAGKDRSRVFI--VG 314
DB 366 NLFETDP-----ARLGELLSTKIVKCAKGLGFTLIGDSSKDEFTIQKS 413
OY 315 IDPNCAAGKGRQLADELLEINGQILYGRSHQNASIIKCAPSKVKIIFIRNKDANO- 373
DB 414 VLSGPPAANGVLRSGDILVRNGRLLGATQKEACDVVAIP-----VNEA 460
OY 374 --MAVCPGNAVEPLPSNSNLONKEPEPTVTSDAVDSLSSKNVQHLPRDQGLGIA 431
DB 461 VDIQVCRGELFIDPAN-----RIVENYAAAKS--RDLEHIDIFKSGSEGFET 508
OY 432 ISEEDTLGVIITKSTEHGVAATDRKLVGDQILAVDEIVVGYPIEKFISLTKAKTV 491
DB 509 IA--DNINGQRKIL--FPSCPNLMGDDIVELDGNVAPRPHQVLMLEKREPIY 563
OY 492 --KLTIHAENPDQAVPSAAGAAGEKKNSSQSLMP-----QSGSEPESEIR 537
DB 564 RCKLVVKKRSPKTRSRTPSAFAYCEPOTNMDSAPLVRSKTPAERQTSITEEQNVR 623
OY 538 NTSRSSTAFASDPATCIIIGCEETIEISKGRGLGSLIYSGSPITLIGATIIHEVYE 597
DB 624 NT-----LORQPAVYSEMEGMSAIPASRRP-----STTLGF----- 657
OY 598 GAACKDRLMAGDQILEVNGIDLRRATNDEAINVLRQTPQRYRLTYRDEAPKKEEVED 657
DB 658 -----ATPNYIPLSGY-----NQKPSD 674
OY 658 TLTIELQKPKGGLGSLIV--GKRNDGVFVSDIVKGIADPDGRLIQDQILLVNGEDVR 716
DB 675 LITVSLIRKP--VGFGFRLLGVESEKPLSVGOIVIGGAEEGRLECGDEIVEIDGHANE 733
OY 717 NASQEVALLKCSLGT--VLEVGRIKAGPFHSEKRPSTQSVSGSISSTFFPLSSGS 774
DB 734 GASHBAVVLLEDAONKVKILV-----RRSRTPARRKSLNS-----AGPS 777
OY 775 TSESLESSSKNALASELOGRLTVEMKKGPTDSLGISIAGV---GSPLDVPPIFIAMH 831
DB 778 -----GSYDVLHHRNENGFGFVLMSSQHKNGSTVGOI-----Q 811
OY 832 PRGVAAGQOKLAVDRIVITIGCTSTEGMTHTQAVNLKNAAGSIEMQV---VAGDSV 887
DB 812 PPSPARCGRSLSVGRVIVANGIDILSLHPTISLIKSGLSVRLTIAPMTAGFVLPM 871
OY 888 VT-----GHHQ-----EPASSLSLFTGISISIQDGLG--PRQCK 921
DB 872 VATLGRNTMNGHTESNGLPRPPPSVEKHPPSYLAFLDGLS---INDRSMNGNLI 927
OY 922 STLESGPDGLFSIYGVGSPHGLPIYVKTFFAKAASEDGRLRGDQIIAIVNGOSLE 981
DB 928 DVTLEKGTGFSGFSTIRG--GQFEGSMPLFVLRIADDGPAKADGRQLVYGQQLTTINGSK 985
OY 982 GVTHEEAVAILKRTKGTVTLMVL 1004
DB 986 GMSHDAIRIKQ--HTMVLVLVL 1007

```

RESULT 7  
571625  
protein-tyrosine-phosphatase (EC 3.1.3.48) RIP - mouse  
N:Alternate names: epidermal growth factor-binding protein; serine proteinase

C:Species: Mus musculus (house mouse)

C:Date: 27-Nov-1997 #sequence\_revision 12-Dec-1997 #text\_change 20-Jun-2000

C:Accession: S71625; S67987

R:Chida, D.; Kume, T.; Mukoyama, Y.; Tabata, S.; Nomura, N.; Thomas, M.L.; Watanabe, F.E.S. Lett. 358, 233-239, 1995

A:Title: Characterization of a protein tyrosine phosphatase (RIP) expressed at a very

A:Reference number: S71625; MUID:95145716

A:Accession: S71625

A:Molecule type: mRNA

A:Residues: 1-2450 <CHI>

A:Cross-references: EXBL.D83966; NID:g1232103; PIDN:BA12158.1; PID:g1232104

A:Experimental source: strain DBA/2; cell line MEL 745A

R:Wolf, B.B.; Brown, M.D.

F.E.S. Lett. 376, 177-180, 1995

A:Title: Epidermal growth factor-binding protein activates soluble and receptor-bound

A:Reference number: S67987; MUID:96105375

A:Accession: S67987

A:Molecule type: protein

A:Residues: 1098-1102 <WOL>

A:Experimental source: submaxillary glands

C:Genetics:

A:Gene: Ptpn13

A:Map position: 5

C:Superfamily: protein-tyrosine-phosphatase, nonreceptor type 13; GLGF domain homolog

C:Keywords: phosphoprotein; phosphoric monoester hydrolase; transmembrane protein; ty

F:566-860/Domain: protein 4.1 membrane-binding domain homology <B41>

F:1089-1165/Domain: GLGF domain homology <GLG1>

F:1361-1437/Domain: GLGF domain homology <GLG2>

F:1495-1574/Domain: GLGF domain homology <GLG3>

F:1769-1840/Domain: GLGF domain homology <GLG4>

F:1863-1937/Domain: GLGF domain homology <GLG5>

F:2203-2422/Domain: protein-tyrosine-phosphatase homology <PPT1>

F:2374/Active site: Cys (phosphocysteine intermediate) #status predicted

F:2380/Binding site: substrate phosphate (Arg) #status predicted

Query Match 10.3%; Score 523; DB 2; Length 2450;  
Best Local Similarity 23.7%; Pred. No. 2.1e-21;  
Matches 240; Conservative 134; Mismatches 346; Indels 294; Gaps 32;

```

OY 196 LEDSPAGKNGTLKPDRIYEAPOSSEPEKAPLCVPPPPSAPFAEMGSDHTQSSAKI 255
DB 1012 LNSKSVASLNSPERRNIESDSSTE--DPQAVYVGMSP-----SSGKS 1055
OY 256 SODVDEDEFGYSKMNIRRYCTLIG--ELMIELEK--HSGLGSLAGNDRSR--MS 309
DB 1056 SSQVPPKDN-----DTLHRKMSIVSSPERETLVNLKDKPKHGLGQIIGKMGCLDG 1110
OY 310 VFIYVGDIPGAGKDRQLQIADELLEINGQILYGRSHQNASIIKCAPSKVKIIFIRNKD 369
DB 1111 VFISAVTPGPRADLDGCLKPGDRLLSVNSVLEGVSHNAAYDILQNAPEADYLVISQPRE 1170
OY 370 AVNQMAVCPGNAVEPLPSNSNLONKEPEPTVTSDAVDSLSSKNVQHLPRDQGLG 429
DB 1171 KPSP-----VPSYVAFANGMKSYTKPAYMDSADPSE-----DQPMFRGTLR 1215
OY 430 IAISEEDTLGVIITKSTEHGVAATDRKLVGDQILAVDEIVVGYPIEKFISL----- 484
DB 1216 HTPESPFGSLG---GLRGSSLSQDSR---TESASLSQSOVNGF---FASHLGRDGM 1263
OY 485 -----RTAKMTVKLTIHAENPDQAVPSAAG-----AASGEKNSS--QSLMPQ 527
DB 1264 QEPQHSPPSPVTTKYNETFSDSNRSAKRRGISDLIELHLCABSDKRDSTYTSQDHQ 1323
OY 528 SGPSEPESEIRNTS--RSSYPAIFASDPATCPIIGCEETIEISKGRGLGSLIYSGSDTL 586
DB 1324 TSKQEPSSSLSTSNKTSFTSSASPK-----RGDTFVELAKTKGSLGSLIYSGVNTSV 1378
OY 587 --GAFIIEVEEGAACKDRLMAGDQILEVNGIDLRRATNDEAINVLRQTPQRYRLTY 644
DB 1379 RHGGIYVKAIIIPKGAASEDGRIVKAGDRVLAVGVSLGEGATTHQAVETLANTQOVHLLLE 1438
OY 645 RDEAPYKKEE-----VCDTLTIELQK--KPGK 669

```

```

Db 1439 KGQVPTSHEDPAGPQSPPPDQDAORQAPKAKOTPHVKDYSPVTEQNTPEVKLFKNSS 1498
QY 670 GLGLS-----YGRKNDTVGVSDIYKGIADPDGLIGDDOILLVNGEDVRNASEA 722
Db 1499 GLGFSFSEHDNLLEPQINGSLVYRKLLFPGQPAESKIDVGLYKNGAPLGLSQDD 1558
QY 723 VAAALKSGISGTVTELEVGRKAKP-----FHSRRPSQTSQVSEGL 763
Db 1559 VISALRGAPVPSLLCLCPAGVLEPIDTFLNLYSPANSFLNLSKRTSQSPSSVEGA 1618
QY 764 SSFFPPPLSG-----STSES----- 778
Db 1619 SSHNGVSGKTKNHCRAFPSRESYSDSHSEGEDSVAPAKMPVTVRYAAEPHEAPRSOE 1678
QY 779 -----LESSK----- 784
Db 1679 ESICAMFTLPKRTKLEKSESHPPPLDVSFGQTCOPPAECAPSDANGKHETHLASOLSK 1738
QY 785 -----KNALA-----SEIQGLRTVEKKGPTSLGISIAG--VSPPLGDVPFIAM 829
Db 1739 EENITTLKNDGNHLEDESELEVLLITLVKSEKSGISGTVTKGSQSISGCVYHDV-----I 1793
QY 830 MHPGVAAQOTKLVGRIRIVYICGSTEGRMTHQAVNLKNAAGSIEMQVAVAGDVSYYT 889
Db 1794 QDP---AKGDRRLKAGRLKIVNDVTYNTMTDAVNLLRAAPKTVRL-----VL 1840
QY 890 GHQEPASSLSFTGLTSTIFODDLGPPQCKSTILEKPGDGLGFSIYGVGSPHGLPI 949
Db 1841 GRILE-----LPRMPYF-----PHLLPDITVTCHEGSELGFSLSGGGSGPHG--VV 1883
QY 950 YKTVFAKGAASEDGRKRGDQIIIAVNGQSLGVTHEBAVALIKRTNGVTLMV 1003
Db 1884 YISINPRSAAVAGDSQIOLDIIVYNGVSTQGMTEEDANRALDLSPLSVLKV 1937

RESULT 8
167629
protein tyrosine phosphatase (PTP-BAS, type 2) - human
C:Species: Homo sapiens (man)
C:Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 16-Jul-1999
C:Accession: I67629
R:Meekawa, K.; Imagawa, N.; Negamatsu, M.; Harada, S.
FEBS Lett. 337, 200-206, 1994
A:Title: Molecular cloning of a novel protein-tyrosine phosphatase containing a membrane
A:Reference number: I53483; MUID:94116679
A:Accession: I67629
A:Status: preliminary; translated from GB/EMBL/DBD
A:Molecule type: mRNA
A:Residues: 1-2466 <RES>
A:Cross-references: GB:D21210; NID:9452191; PIDN:BA04751.1; PID:9452192
C:Superfamily: protein-tyrosine-phosphatase, nonreceptor type 13; GLGF domain homology:
F:574-668/Domain: protein 4.1 membrane-binding domain homology <B41>
F:1354-1430/Domain: GLGF domain homology <GLGF>
F:2218-2437/Domain: protein-tyrosine-phosphatase homology <PTP>

```

```

Query Match 10.0%; Score 509.5; DB 2; Length 2466;
Best Local Similarity 22.3%; Pred. No. 1.3e-20;
Matches 256; Conservative 146; Mismatches 395; Indels 351; Gaps 41;

```

```

QY 142 ELQNTAVSNMQRVELMREPKSLGISYGGMGSRSLNSG----- 184
Db 868 QLOMRAROSNODADIERASFRSLNLAESVRFNMGRRAISTGSLASTLKLAVRPLSV 927
QY 185 --EVMRGIFIKHY-----LESDPAKNGTL-----KPGD--RIVAPQSESEPEKAPLCS 231
Db 928 QAEILTKRISCESELVLPYLPLOASKKNDKASWEKPREMSKSYHDLSOASLYPRKKNIV 987
QY 232 VPPPPSAFAEM-GSDHTOSSAS-----KISQVDKED 263
Db 988 NMEPPPTVAELVGRKPHQMSRDAESLAGVTKLNNSKSVASLNRSPERRKHEDSSSIE 1047

```

```

QY 264 EFG--YSWKNIERYGTLTG---ELHMLELEK-GHSGIGLSIAGNKDRSR--MSGVTVGI 315
Db 1048 DRGQAVLDVHKRMMSYSSPEREITLVNKLKAKKRGIGLGPQIIGEKMKRGLDGLFISV 1107
QY 316 DPNAGKDGRLQIADLELLEINGQILYGRSHQNASIHKCAPSVKIIETRNKDAVQMA 375
Db 1108 APGCPADLDGCLKPGDRILISVNSVSLGSHHAILEIQNAPEDVTVLISQPKKISKVP 1167
QY 376 VCPGNAVEPLPSNSENQNKTEFEPTVTTSDAVALDSFEKN-----VOHLELP--DOGG 427
Db 1168 STPVH---LTNEMKNMKKSS---YMDSADISSKDHMSGTLRHISNSFGSPGG 1219
QY 428 L'-GIAISEEDTLGVIILKSLTEHGAATDGRKLVGQDIIAVALDEIYVGYPIEKFIILK 485
Db 1220 LREGSLSSQDSRTEASLSQGVNGFPAS---HLGQGTW---QESQHGSPSPVIS--- 1269
QY 486 TAKMTVKLTIAENPDSONVPSAAGASGEK-----NSQSILMPQSGSPES 535
Db 1270 --KATEKETFTDSNQSTKRGISDVYDSDRGSDDEATYSSSDHQHPKQ---ESSS 1324
QY 536 IRRNTSRSTPAIFADPATCPIIPGCEETIEISKRGTLGSLIVGSGDTLL--GAFIIE 593
Db 1325 SVNTSNKNMFTFESSPEK---PGDIEVELANDNSLIGSVTGAVNTSVRHGCIYKA 1380
QY 594 YVEBGAACKDGRILWAGDQIILEVNCIDLRKATHDAIVNLROTQVRRLTYRDEAPYKEE 653
Db 1381 VIFQGAESDGRIRHKGRVLAVNGVSLGATHKQAVETLRNTGQVHLLLEKQSPSKKE 1440
QY 654 EV---C-----DLTIELOKPKKGLS----- 674
Db 1441 HVPVTPQCTLSDNAQAGPEKVKTKTQVKDYSFVTEENFEVLEFN--SSGLGFSFSRE 1499
QY 675 ---IVGRNDTVGVSDIYKGIADPDGRILQDQIILLVNGEDVRNASEAVALLCSTL 731
Db 1500 DNLIPEQINASTYVRKTLFPGQPAESGKIDYGVILKVNAGSLKGLSQGEVIALNGTA 1559
QY 732 GTVTVLEVGRKAKP-----PHNSE----- 749
Db 1560 PEVFLLCRPPPGVLEPIDTALLTPLOSPAVQVLPNSSKDSQSPCVBQSTSDENEMSK 1619
QY 750 ---RRPSQTSQVSEGS-----LSGFTF-----PL 770
Db 1620 SKKCKSPSRKDSYSDSSGSGEDLVTAIPANISNISTWSALHQTLSNMVSOAQSHHEAPK 1679
QY 771 SGSST-----SESLSSSK----- 784
Db 1680 SQEDTICTMFFYPQKIKNKPEFEDSNPSLPDPMAPEGSYQPGESASSSMOKYHITHI 1739
QY 785 -----KNALASEIQGLR-----IYEMKKGPTSLGISIAG--VSPPLGDVP 824
Db 1740 SEPTROENMTPLKNDLENHLEDEFELEVEILLITLKSEKSGISGFTVTKGNORIGCVYHDV- 1798
QY 825 IFIMMHPGTVAAGTOKLRVGRDRIYVTCGSTEGRMTHQAVNLKNAAGSIEMQVAVAGD 884
Db 1799 ---IQDP---AKSDGRILKPEDRLIKVNDVTYNTMTDAVNLLRAASKTVRL----- 1844
QY 885 VSVVTGHHOE-PASSLSFTGLTSTIFODDLGPPQCKSTILEKPGDGLGFSIYGVGSP 943
Db 1845 ---VIGVLELPRPML-----PHLLPDITLVNCKRELGFSLSGGHDSL 1885
QY 944 HGDLPYKTVFAKGAASEDGRKRGDQIIIAVNGQSLGVTHEA-----VALIKR 994
Db 1886 Y--OVVYISINPRSAVIAIEGQLDLIVHYVNGVSTQGMTELEVNRALDMSPLSLVKA 1943
QY 995 TKGTVTLM 1002
Db 1944 TRNDLPVY 1951

```

```

RESULT 9
167630
protein tyrosine phosphatase (PTP-BAS, type 3) - human
C:Species: Homo sapiens (man)

```



|                       |                   |                    |             |              |
|-----------------------|-------------------|--------------------|-------------|--------------|
| Query Match           | 9.9%;             | Score 501.5;       | DB 2.       | Length 2294; |
| Best Local Similarity | 22.8%;            | Pred. No. 3.2e-20; |             |              |
| Matches 263;          | Conservative 142; | Mismatches 400;    | Indels 351; | Gaps 44      |

|    |      |   |      |
|----|------|---|------|
| Qy | 750  | -----RRPQSTQSVSESS-----LSSFP-----                             | 768  |
| Db | 1440 | DENEMSKSKKOCKSPSRSDYSODSSGSGEDDLVTAPANISNSTWSALHOTLSNMWSOA    | 1499 |
| Qy | 769  | -----PLSGSSST-----SESLESSSK                                   | 784  |
| Db | 1500 | QSHHEAPRSQEDTICTMFTYPOKIPNKEPFDNSNPRLPPDMAPQSYQPOPSSEASSSSSM  | 1559 |
| Qy | 785  | -----KNALASEIQGLR-----TWENKKGPTDSLGISIAG---V                  | 816  |
| Db | 1560 | DKYHHIHISEPTROENMTPLKNDLENHLEDELEVEELLITLIKSEKSGSLGFTVTKGNORI | 1619 |
| Qy | 817  | GSPGLGDVPIFIAMMHPRTVAQAOTOKLRVGDPIVITICGSTECSMPTOAVNLKNAAGSIE | 876  |
| Db | 1620 | GCYHADV-----IQDP---AKSDGRLKPGDRLIKVNDTDTNMHTDVAVNLRAASKTVR    | 1671 |
| Qy | 877  | MOVVAGGDVSVYTGHHOE-PASSSLSTFGTLSTSIFFODLGPPOCKSTTLERGPDLGFS   | 935  |
| Db | 1672 | L-----VIGRVLELPRIPML-----PHILPDTTLICNKEELGFS                  | 1705 |
| Qy | 936  | IVGTVGSPGHGLPIYVKTVFAKGAASEDGRLKRGDQIIITAVNGSLEGTTHEDA-----   | 988  |
| Db | 1706 | LCGGHDSLEY-QVYVYISIDINRSVAIEGNQLDLVDIHYVGVSGTQMTLEEVNRALDMS   | 1763 |
| Qy | 989  | --VALKRTKGTVTLM   | 1002 |
| Db | 1764 | LPSTVLKATRNDLPVY  | 1779 |

RESULT 10  
A:Accession: A54971  
protein-tyrosine-phosphatase (EC 3.1.3.48), nonreceptor type 13 - human  
N:Alternate names: Fas-associated phosphatase FAP-1; protein-tyrosine-phosphatase hpp1  
C:Species: Homo sapiens (man)  
C:Date: 11-Nov-1994 #sequence,revision 08-Feb-1996 #text,change 01-Dec-2000  
C:Accession: A54971; A55114; 159595; 153483; 546955  
R:Barville, D.; Ahmad, S.; Stocco, R.; Shen, S.H.  
J. Biol. Chem. 269, 22320-22327, 1994  
A>Title: A novel protein-tyrosine phosphatase with homology to both the cytoskeletal  
A:Reference number: A54971; MUID:94350988  
A:Accession: A54971  
A:Molecule type: mRNA  
A:Residues: 1-2490 <BRAN>  
A:Cross-references: GB:U12128  
A:Note: sequence shown follows authors' translation at positions 62-63  
R:Stas, J.; Claesson-Welsh, L.; Heldin, C.H.; Gonsky, L.J.  
J. Biol. Chem. 269, 24087-24089, 1994  
A>Title: Cloning and characterization of PTP11, a protein tyrosine phosphatase with s  
A:Reference number: A55114; MUID:95014139  
A:Accession: A55114  
A:Molecule type: mRNA  
A:Residues: 1-61, 'GS', '64-839', 'D', '841-1055', '1075-1133', 'FH', '1136-1210', 'I', '1212-1383', '1386  
A:Cross-references: GB:X80289; NID:9515030; PIDN:CAA56563.1; PID:9515031  
R:Sato, T.; Irie, S.; Kitada, S.; Reed, J.C.  
Science 268, 411-415, 1995  
A>Title: FAP-1: a protein tyrosine phosphatase that associates with Fas.  
A:Reference number: 159595; MUID:95232528  
A:Accession: 159595  
A>Status: translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1279-1888 <RES>  
A:Cross-references: GB:U34583; NID:9806291; PIDN:MAC41755.1; PID:9806292  
R:Maekawa, K.; Imagawa, N.; Nagamatsu, M.; Harada, S.  
FEBS Lett. 337, 200-206, 1994  
A>Title: Molecular cloning of a novel protein-tyrosine phosphatase containing a membra  
A:Reference number: 153483; MUID:94116679  
A:Accession: 153483  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-61, 'GS', '64-839', 'D', '841-1210', 'I', '1212-1383', '1389-2299', 'OM', '2302-2490 <REZ  
A:Cross-references: GB:D21209; NID:9452189; PIDN:BA004750.1; PID:9452190  
C:Genetics:  
A:Gene: GDB:PTPNI13

A:Cross-references: GDB:306348; OMIM:600267  
 A:Map position: 4q21.3-4q21.3  
 C:Superfamily: protein-tyrosine-phosphatase, nonreceptor type 13; GLGF domain homology;  
 C:Keywords: alternative splicing; phosphoprotein; phosphoric monoester hydrolase; tyrosi  
 F:574-868/Domain: protein 4.1 membrane-binding domain homology <B41>  
 F:1089-1175/Domain: GLGF domain homology <GLG1>  
 F:1373-1434/Domain: GLGF domain homology <GLG2>  
 F:1511-1590/Domain: GLGF domain homology <GLG3>  
 F:1799-1870/Domain: GLGF domain homology <GLG4>  
 F:1893-1967/Domain: GLGF domain homology <GLG5>  
 F:2242-2461/Domain: protein-tyrosine-phosphatase homology <PPT1>  
 F:2413/Active site: Cys (phosphocysteine intermediate) #status predicted  
 F:2419/Binding site: substrate phosphate (Arg) #status predicted

Query Match 9.7% Score 494.5; DB 1; Length 2490;  
 Best Local Similarity 22.6% Pred. No. 9.2e-20;  
 Matches 243; Conservative 140; Mismatches 371; Indels 319; Gaps 38;

QY 162 EPRKSLGISIVGRCMSRLSNGEVMRIGITKHVLEDSPPACKNGTLKPGDRIVEAPSOSE 221  
 DB 990 EPPQTVAEVLGKPSHQMSRSDAELAGV---TKLNNSKVASLNRSPERRKHSSDSSI 1046  
 QY 222 SEPEKAPLCSPPPPPSAFAEMGSDHTOSSASKISQDYDKEDERCYSKKNIREKIGITG 281  
 DB 1047 EDPGQAVYLGMT-----MHSSGNSSQVPLKENDV-----LHKRWSTVSS 1086  
 QY 282 ---ELHMLEER-GHSGIGLSLAGNKRDR--MSVFIQIDPNCAGAKGRQIADILLE 335  
 DB 1087 PERETTLVNLKDAKYGFGFDIGEEKMGRDLGIFISSVAPGAPADLDGLCKPDRILIS 1146  
 QY 336 INGQLYRSQNMASIIKCAKPSKVKITFIIRKDAVNOMAVPCGNAPVPLPSNSENLOMK 395  
 DB 1147 VNSVSLGVSHHAAIEIIONPEDVTLTISQPKKISKVSPSTPVH---LNNEMKNMYKK 1202  
 QY 396 ETEPTVTSDAAVDSLFKN-----VOHLELPK--DQGL--GTAISEEDTLGCVITKS 445  
 DB 1203 SS---YMODSAMOSSKSDHMSKGTLLHISENSRPGSGIGREGSLSDSRTESASISQ 1238  
 QY 446 LTHGVNATDRLKVGDDIILAVDEIVVGPRIEFISILTKAKMTVKTLIAENPDQAV 505  
 DB 1259 SQVNGFFAS---HLGDQTV---QESQHGSPSPVIS-----KATEKETFDSDNSQSKTK 1306  
 QY 506 PSAGAASGEKK-----NSSQSLMPQSGSPPESTIRNTRSRSTPAIRFASDPATC 555  
 DB 1307 PGISDVTDYSDRGSDMDDEATYSSDQHOTPKQ--ESSSSVNTSNKKNFTKFFSPPK- 1362  
 QY 556 PIRGCETTIEISKRTGLGSIY-----GGSDPLL--GAFIHEVYEEGAACKDGRIMA 608  
 DB 1363 ---FODIFEVLELAKNDNSLGISIVVLPDKGVNTSVRHGITYKAVIPOGAESDGRILK 1419  
 QY 609 GDOILEVNGIDLKRAITHDEAINVLKQTPQVRVLTLYRDEAFYKEEV-----C----- 656  
 DB 1420 GDRVLAVNGVSLGEGATHKQAVETLRNTGOVYHLLLEKQSPTSKEHVVTPOCTLSDQNA 1479  
 QY 657 -----DTLTLELKKRGKGLGLS-----YVKKRNDTGVFV 686  
 DB 1480 QGQGEPEKVKTTQVKNYSFVTEENTFEYKLEKN--SSGLGFSPSRBDNIIPQINASTIRV 1538  
 QY 687 SDIYKGIADPDRLIGDQDILLVNGEDVNRASQAVALLKCSIGVTLTEVGRIRKAG-- 744  
 DB 1539 KKLPPGQPAASGKIDGVILKYNASLKLGLSQOEVLSALRGTAPEVFLLCRPPGCVL 1598  
 QY 745 -----PRHSE-----RRPSQTSQVS 759  
 DB 1599 PEIDTALLFPQSPAQVLPNSSKDSQPSQVEQSTSSDENEMSKSKQCKSPSRSDYS 1658  
 QY 760 EGS-----LSSFTF-----PLSGSST----- 775  
 DB 1659 DSSGSGEDDLVTAPANISNTWSALHQTLSNMVSOASHHAPKSOEDTICTWFFYPOK 1718  
 QY 776 -----SESLSSSK-----KNA 787

DB 1719 IPNKEFEDSNPSPLPPDMABGQSYQPOSASSSMDKVHIIHISEPTROENMTPKND 1778  
 QY 788 LASEIOLGR-----TVEKKKPTDSLGISING---VSGPAGDVPFIAMHPGVAAQT 839  
 DB 1779 LKNHDEFELEVELLITLIKESKSLGFTYKGMORIGCYHVD-----IODP--AKSD 1830  
 QY 840 QKLRYGRIYVTCSTEGMTHTQAVNLKNASGSIEMQVAGGDVSVYGNHDE-PASS 898  
 DB 1831 GLRKPGRDLIVNDPDTNMHTDAVNLRAASKTVRL-----VIGRVLELPRIIP 1880  
 QY 899 SLSTGTLSTIFQDDGLGPPCKSTTLERPGDGFSGFISVGGYSGPBHGLPIYKTVAKG 958  
 DB 1881 ML-----PHLLPDTITCNKEKEELGFFSLCGCHDSLY--QVYIYSDINPRS 1922  
 QY 959 AASEDRGLKRGDDIILAVNGSLEGVTHEE-----VALKRTGVTIM 1002  
 DB 1923 VAAIEGNQLQDLVHYVNGVSTQGTLEEVNRALDMSLPVLAKATKNDLPV 1975

## RESULT 11

JE0209  
 brain-specific angiogenesis inhibitor-associated protein 1 - human  
 N:Alternate names: BAI1-associated protein 1; BAP1 [misanomer]  
 C:Species: Homo sapiens (man)  
 C>Date: 21-Aug-1998 #sequence, revision 21-Aug-1998 #text, change 18-Aug-2000  
 C:Accession: JE0209  
 R:Shiatsuchi, T.; Futamura, M.; Oda, K.; Nishimori, H.; Nakamura, Y.; Tokino, T.  
 Biochem. Biophys. Res. Commun. 247, 597-604, 1998  
 A:Title: Cloning and characterization of BAI1-associated protein 1: A Pdz domain-conta  
 A:Reference number: JE0209; MUID:58321173  
 A:Accession: JE0209  
 A:Molecule type: mRNA  
 A:Residues: 1-1256 <SH1>  
 A:Cross-references: GB:AB010894; NID:93370997  
 A:Experimental source: brain  
 C:Genetics:  
 A:Gene: GDB:BAIAP1; BAP1  
 A:Cross-references: GDB:9864783  
 A:Map position: 3p14.1-3p14.1  
 C:Superfamily: WW repeat homology  
 F:300-337/Domain: WW repeat homology <WW1>  
 F:359-396/Domain: WW repeat homology <WW2>

Query Match 8.9% Score 453.5; DB 2; Length 1256;  
 Best Local Similarity 22.4% Pred. No. 7.1e-18;  
 Matches 184; Conservative 126; Mismatches 283; Indels 229; Gaps 30;

QY 378 PGNAVEPLPSNSENLOMKETPTVTTSDAAVDSLFKNVQHLELPKQDGGIGLAI---SE 434  
 DB 442 PSN---PEPAEVPLOG--PFTTRNPSELK---CKFITKLRKSSRGFGFTVVGDE 491  
 QY 435 EDTLSGVTKSLTEHGAATGRKLVGDDIILAVDEIVVGPRIEFISILTKATMTVK-- 492  
 DB 492 PDEF--LQIKSLVLDGRALDGMKMETGDVIVSMDTCVILGHTHAQVYKIFQISITIGASVD 549  
 QY 493 -----LTIHAENPDQAVPSAAGAASGEK---NSSQSLMPVQSGSPPEPSIR--N 538  
 DB 550 LELCRGYPLPDPDPDPTSLVTSVAL--DKEPLIVNGGETYDSPAHSKTKGVNGMK 606  
 QY 539 TSRSTRAIFSDP-----ATCFPIRGCETTIEISKRTGLGSIY---G 580  
 DB 607 DARPSSPADVASNSHGVPNDTVLASSIAQPEL---IIVHIVKGMGFGFTIADSPG 662  
 QY 581 GSDTLLGAFIHEVYEGACACKDGRLMNAGDQILEVNGIDLKRAITHDEAINVLKQTPQVR 640  
 DB 663 GGGQRAQIV-----DSPRCRG--LKEGLDLYEVNKKKNQALTHNOVDMVLVECPRGSE 714  
 QY 641 LTL-----YRD-----EAYKE 652  
 DB 715 VTLVQNGGLVPKPKSPQLEKKDSQNSQSHSVSHRSRLHTASPSSHTQVLPPEFPPE 774  
 QY 653 EEVCDTLTLELKKRG----- 668

Db 775 AQPAPDTSAGGKPPDFKMAOSRWYENRPMSPASGLSKGERERINSTNGECPI 834  
 QY 669 -----KGLSLIVKRNDRG--VFVSDIYKGIADDDGRLIGDQILLVNG 712  
 Db 835 PDYQODIFLWKEKKEFGFRILG-GNEPEPIGHIYVPLGADIDGRSDELDICVDG 893  
 QY 713 EDVRNASQEAVALLK--CSLCTVTLVGR-----IKAGFHSERP--S 753  
 Db 894 TPVICKSHQVYVLMQAAKQGHVNLTVRKVFAVPKTEENVPASSHSHSNPASLT 953  
 QY 754 QTSQVSEGLSFTPLPSLSTSELES--SKKNALASEIQLRTYEMKKGPTDLSIGSI 812  
 Db 954 EKKRTPQSGNSLNTVSSSGSTSGISGGSGSVTVVQPYVEIRRGNEGFGPIY 1013  
 QY 813 AGGVSPGLDVPF---FLAMHPTG-----VAAQTKLRVDRIVTICGTSTEGMHT 862  
 Db 1014 VSSVSRPEAGTTFAGNACVAMPKIGRIEGSPADRCGLKVGDRILAVNGCSITNKS 1073  
 QY 863 QAVNLKMASGSIEMQVYAGD-----VSVYTGHH---QEPASSLSFTGLTS 907  
 Db 1074 DIVNLKIKAGNVTLRILIPGDESSNATLLTNAEKIATITTTTPSOQGTQETRNTPK 1133  
 QY 908 TSIFQDDLCPPCK-----SITLERGPDGLFSIVGSPHGDLPYVKTVPKGAAS 961  
 Db 1134 ESQF--EFKAPQATQODEFYTELERGAKGFGFSLRG--REYMDLIVRLAEDGP 1188  
 QY 962 EDRLKRGDQIIAVNGSLGVTHHEAVAILKRTKGTVTLMV 1003  
 Db 1189 RCGMKRIGDEILEINGETTKNMKHSRALEILKNGGRVRLFL 1230

## RESULT 12

channel associated protein of synapse 2 - rat  
 C:Species: Rattus norvegicus (Norway rat)  
 C:Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 21-Jan-2000  
 C:Accession: T10811  
 R:Ririe, M.; Hata, Y.; Takai, Y.  
 submitted to the EMBL Data Library, April 1996  
 A:Description: Cloning of new isoforms of PSD-95/SAP90 related genes.  
 A:Reference number: 217166  
 A:Accession: T10811  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-852 <IRT>  
 A:Cross-references: EMBL:U53368; NID:G1517939; PID:G1517940  
 C:Genetics:  
 A:Note: PSD-95/SAP90-related gene 2 (chapsyn isoform2)  
 C:Superfamily: discs-large tumor suppressor; GIGF domain homology; guanylate kinase hom  
 F:198-226/Domain: GIGF domain homology <GIG>  
 F:343-601/Domain: SH3 homology <SH3>  
 F:663-840/Domain: guanylate kinase homology <GKI>

Query Match 8.0%; Score 407.5; DB 2; Length 852;  
 Best Local Similarity 28.4%; Pred. No. 1.7e-15;  
 Matches 153; Conservative 68; Mismatches 212; Indels 111; Gaps 22;

QY 498 ENPDQAAPSAGAASG-----EKKNS-----QSLMPQSGSEPESTIRTSR 541  
 Db 22 DGFPHSHLPRLTHEVRGPELVHSEKNSQIENHGVYVLSHLSPLKSPAP--IIVNDT 80  
 QY 542 STPAIFASDPATCPITIGCEETIE-----ISKRTGLGLSTVGSST--LLG--AFI 592  
 Db 81 LD-----TIYVNGTEIEFEFEETTLERNGSLGFSIAGGTDNHIDDDGIFIT 130  
 QY 593 EYEEGAACKDGRMLAGDOILEVNGIDLRKATHDEAINVLRQTPQVRVRLTLVRDEAPYKE 652  
 Db 131 KIIPEGAAEDRLKVNDCILRVNEDVSEVSHSKAVLEAKGASTIVRLY--RRRRPIL 189  
 QY 653 EEVCDTLTLELQKPKGKGLSIVCK-----RNDTGVFVSDIVKGLIADPDGRLIGDQ 706

Db 190 TVV-----EIKLFKGP-KGLGFSIAGVGNQHIQDINSIYVTTIIDGAAQKGRLOVGR 244  
 QY 707 ILLVNGEDVRNASQEAVALLKCSIGTVTLVEXGRKA-----GP---FHSERPQTS 756  
 Db 245 LLMVNNYSLEEYTHEEAVAILKNTDYYLKVKGKPTTIYMDPYGPPDITHSYSPMENH 304  
 QY 757 QVSEGLSFPPLPSLSTSELES--SKKNALASEIQLRTYEMKKGPTDLSIGSIAGV 816  
 Db 305 LLS-GNNGTLEKTSPLPSISGERSYPIPKHMLV--EDETTRPE---PYSTVKNLCKDP 358  
 QY 817 GSPGLDVPFIIFAMHPTGVAQTKLRVDRIVTICGTSTEGMHTQAVNLKMASGSI 876  
 Db 359 ASPRHYSPV-----ECDKSF-----LLSTPYPHN 383  
 QY 877 MOVVAGDVSVYTGHHQEPASSLSFTGLTSTIFQDDLCPPCKSITLERGPDGLFSI 936  
 Db 384 LCLLPDSMTSHSQSTATRQPSVTLQRAISLE-----GEP--RKVVLHKSGTGLGFINI 435  
 QY 937 VGVYSPHGDLPYVKTVPKGAASEDRKRGDQIIAVNGSLGVTHHEAVAILKRTK 996  
 Db 436 VGG---EDGE-GIFVSLIAGPADLSELGRDQILSVNGIDLKGAHQAAALAKGAG 491  
 QY 997 GTVTLM 1002  
 Db 492 QTVTII 497

## RESULT 13

hypochemical protein C25F6.2 - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C:Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 20-Sep-1999  
 C:Accession: T15617  
 R:Bentley, D.  
 submitted to the EMBL Data Library, October 1995  
 A:Description: The sequence of C. elegans cosmid C25F6.  
 A:Reference number: 218377  
 A:Accession: T15617  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-1131 <BEN>  
 A:Cross-references: EMBL:U39742; NID:G1049455; PID:G1049459; PIDN:AAA80434.1; CESP:C2  
 A:Gene: CESP:C25F6.2  
 A:Introns: 105/1; 210/3; 283/3; 316/1; 346/2; 463/3; 566/1; 722/2; 897/1; 991/2; 1032

Query Match 7.9%; Score 399.5; DB 2; Length 1131;  
 Best Local Similarity 24.4%; Pred. No. 7.4e-15;  
 Matches 185; Conservative 102; Mismatches 322; Indels 149; Gaps 27;

QY 126 DIPELPEREGGEES---ELQNTAYSNMNPRAVELMRPESKLSISIVGKMGSR 181  
 Db 171 EVRRIYERLEG-GPHSYNSRPATATSTSNVNLSTPLISDLDRDGRGFSYLGNGLG 229  
 QY 182 SNGEVMKRGFTKHYVEDSPAGKNGTLKPODRIVEAPSOSESPEKAPLCSVPPPPSAFA 241  
 Db 220 GNG-----LQNGLLSPYNSST---HYLHERQRTSHQGTWRETTTTRVDP 278  
 QY 242 EMGSDHTQASAKISQVDKDEDFYSWMKNIRERGTLTGELHMLELEKSGSLSLAG 301  
 Db 279 RRVVEHT-----GVDDHGRKM-----ELENYLLEKRGHGLGFSITG 315  
 QY 302 NKDR---SRMSVETVIGIDPNCAGKGRGLQIADLELINCQILVGRSHQNASIICAP 357  
 Db 316 GMDQPTEDGDTTIVYTNIEEGALADGMRKNDITFANNNTNCENKHEVAVNALKSSG 375  
 QY 358 SKVKIIFIRKND-----AVNMAVCP-----GNAVEPLPSNSELQKKEEP 399  
 Db 376 NVVYSLSLRKKEAPLPIGNGFGSTYLRSGVTYSVAGNLQHAHISPSAPAIHPPPPP 435  
 QY 400 TWTSDAAVDSLSEK---NQVHLELPKDOGLGIAIS-----EDDTLSGVLIKSLT 447

Db 436 VHHGSLQSLVGYQSRPRNPTSVIDLKVGANGLGEISLAGOGNEHVAGDT--DIYVTKII 493  
 448 EHGVAATDGRKLVGDOITLAVDEIVGVPIKFIISLTKAKMTATLHA-----ENP 500  
 494 EHGAAEEDGRVRVDKILEVDHSLINTHENAVNLKNTGNRRRLILQOSTGALFNDSA 553  
 501 DQAQVPSAAGAASEKKNSQSLSLWVPS---GSPEPSIRNTSRSSSTPAIFASDPATCP 556  
 554 SQQFMPTPIILPSSVODYNRSQMSQSHLSYGGP-----LNTSYSS-----CAPAIIP 602  
 557 IIPGCEETIEISKRTGLGSLVSDTLGAFIHEHYEGACAKGRMLAGQOILLEV 616  
 603 LEP---RPVQLVKGQNGENGVIGEDV--EPIYISFVLPGVADLSGNVATGVLELVN 657  
 617 GIDLRKATHDEAINVLKOTPORVRLTLYRDEAPYKEEVDLTILEKKPGKGLGSI 675  
 658 GVLVRLNATPKRAALRNAGNPNVLTLYQYRQ---EQIPIES---KIEKARNVIAQSR 710  
 676 VGRKNDTGVFVSDIVKGIADPGR-----LIQGOIILLVNGED-----VRN 717  
 711 MGLSRKSEVYRALFD---YDPSRENSVAPHRSMGFNYGDIILHINSDDDEMTARKVHE 767  
 718 ASQEAVALKLCSLGTVLEGRKAGPNS-----ERRPSOTSQVSESSL 763  
 768 NGEETAGCVLEPSKRVKREKRLKRYOVNFNSGSLGRNSSTGLNKRGRSRLQSES-- 825  
 764 SFTFPLSGSTSESLSSSKKNALASEIOGLRTVEMK 801  
 826 RKPFVAKSTDRNLNLEESS--NVAEEVMSYQAVEQO 861

RESULT 14  
 601974  
 channel associated protein of synapse - human

C:Species: Homo sapiens (man)  
 C>Date: 21-Dec-1996 #sequence\_revision 06-Jun-1997 #text\_change 21-Jan-2000  
 C:Accession: G01974  
 R:Kim, E.; Cho, K.  
 submitted to the EMBL Data Library, July 1995  
 A:Reference number: G08966  
 A:Accession: G01974  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-870 <KIM>  
 A:Cross-references: EMBL:U032376; NID:g1463025; PID:g1036790  
 C:Genetics:  
 A:Gene: chapsyn-110  
 C:Superfamily: discs-large tumor suppressor, G1GF domain homology; guanylate kinase homology  
 F:198-276/Domain: G1GF domain homology <G1G2>  
 F:343-601/Domain: SH3 homology <SH3>  
 F:601-858/Domain: guanylate kinase homology <GKI>

Query Match 7.8%; Score 397.5; DB 2; Length 870;  
 Best Local Similarity 28.1%; Pred. No. 6,4e-15;

Matches 154; Conservative 69; Mismatches 205; Indels 121; Gaps 23;

Db 500 PDSQAVPSAAGAAGSG-----EKKNS-----OSLWVPSGSGPEPSIRNTSRSS 543  
 24 PHDSISLRLHREHVPSELVHSEKNLSQIENVHGYVLOSHISPLKASAP--IIVNTDTLD 82  
 83 -----TIPVNGTLEIEEIEETLEKNGSLGFSIAGCTDNDHIGDDPEIFITKI 132  
 544 TPALFADPATCPIIPGCEETIE-----ISKRTGLGSLVSGSDT--LLG---AFIIEV 594  
 595 YEEGAAKDGRLAGDOILEVNGIDLKATHDEAINVLKOTPORVRLTLYRDEAPYKEE 654  
 133 IPGAAAEEDRLRVNDCILIRNEVDVSEVSHSKAVEALKEGSLARLYV--RRRRPILETY 191  
 655 VCDTLTELOKKPGKGLGSLVSK-----RNDTGVFVSDIVKGIADPGRGLQGOIL 708  
 192 V-----EIKLRKGP-KGLGFSIAGVGNQHIPGDNSTIVTKIIDGGAOKDGRLQVGRLL 246

Qy 709 LVNGEDVRNMSQEAVALKLCSLGTVLEVGRIKA-----GP---FHSRRPSQTSQV 758  
 247 MVNNYSLEEVTHEEVAVALINKTSEVVLKGNPTIYMTDYGPRDITLHVSPEMHNLL 306  
 759 SEGSLSFTPLPGLSSSTSESLSSSKKNALASEIOGLRTVEMKKGPIDSLGISTAGVGS 818  
 307 S-GNNGTLEYKTSLPPIISPGRYSPIRKHMVLD--DYTRPPE-----PYSVYVNRKCDRPAS 360  
 819 PLGDPVI-----FIAMHPIGVANQOTOKLRVGRIRIVICGSTEGMTHTQAVNLKNASG 873  
 361 PRHSIPVECKSFL-----LSAPYSHYLG-----LPPDSMTSHSQHSTATRQPSM 407  
 874 SIEMQVAGDVSVYVTHGHOEPASSLSFTGLSTSIPODDLGPQCKSTTLERKPDGLG 933  
 408 TLQRAVSELG-----EP-----RKVLHKGSTGLG 432  
 934 FSTVIGVSGSPHGDLPYVKTYFAKGAASEDGRKRGDOITAVNGQSLGVTHEEVAVALK 993  
 433 FNIVGG---EDGE-GIVFSFLIAGPADLSGELQRGDQILSVNGIDLKGASHQEAALAK 488  
 994 RTKGTVTLM 1002  
 489 GAGOTVITII 497

RESULT 15  
 T14152

synaptic scaffolding protein S-SCAM - rat  
 C:Species: Rattus norvegicus (Norway rat)  
 C>Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 18-Aug-2000  
 C:Accession: T14152  
 R:Hi Rao, K.; Hata, Y.; Ide, N.; Takeuchi, M.; Irie, M.; Yao, I.; Deguchi, M.; Toyoda, J. Biol. Chem. 273, 21105-21110, 1998  
 A:Title: A novel multiple PDZ domain-containing molecule interacting with N-methyl-D-  
 A:Reference number: Z17889; NMID:98361985  
 A:Accession: T14152  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-1277 <HIR>  
 A:Cross-references: EMBL:AF034863; NID:g3411052; PID:g3411053; PIDN:AMC31124.1  
 C:Genetics:  
 A:Gene: S-SCAM  
 A:Function:  
 A:Description: may assemble receptors and cell adhesion proteins at synaptic junction  
 C:Superfamily: WW repeat homology  
 F:302-339/Domain: WW repeat homology <WW1>  
 F:348-385/Domain: WW repeat homology <WW2>

Query Match 7.8%; Score 395; DB 2; Length 1277;  
 Best Local Similarity 19.9%; Pred. No. 1.6e-14;

Matches 216; Conservative 140; Mismatches 353; Indels 378; Gaps 39;

Qy 183 NGEVMSGIFTKHVEDSPAGKNGTLKPGDRIVAPSSQSESEPEKAPLCVSP----- 233  
 245 NGVYVTPESSEH--EDKSAGASG-----ETPSSQ--PYAPVYSQPEELKQMD 289  
 234 -----PPPS--AFAMGS-----DHTGSASKIS-----QDVOKED 264  
 290 TKSTKPEENEDSDPLPDNMEMAYTEKGEYFIDHNTKTSWLDRLAKKAPAECKENE 349  
 265 FGYSWKNIRER-YGT----- 278  
 350 LPFGMEKIDPIYGVYVDHINRTQENPVLEAKRKLQOHNPHELTGKPLQAPGFE 409  
 279 -----LTBELMIELEKSGSLGSLAGNKRDSRMSVFIVGIDPGAAGKGRLO 328  
 410 KPLFTRDASOLKGTFLSTTLKSNMGFGFTIIGG--DEPDEFLOVKSVIPGPAQDKME 468  
 329 IADELLEINQOILXGRSHOASSIICAP--SKVKIIFIN-----KQAVNQMANVCPG 379  
 469 TGDVIYINEVGVGHADVAVKLFQSVPIGQSVNLVLCRGYLPDPDEPANSMV--PP 526

```

QY 380 NAV--EPLPSNSENLOKNET-----EPTV 401
| : | | |
Db 527 LAIMERPPVWYNGRHNNTYLETYSRTSQSPDITDRPPHLSHMPADGOLDGTYPPV 586
| : | | |
QY 402 TTSDAVIDLSFRKNOHLEPKDOGLGIAISEDTLSCVITIKSLTEHGVAAATDGRKVG 461
| : | | |
Db 587 HDONYSVASSGATQALMLMTLTIVKAKGFGFTADSPGTGRVKQILD--IQCGP-LCEG 643
| : | | |
QY 462 DQILAVDDIYVGYPIEKTISLKTAKM--TVKLTJHA-----ENPDS 502
| : | | |
Db 644 DLIVEINQONVONLSHTEVVDILKDCPVGETSLIHRGFFSPWKTPrMVDWRENQGS 703
| : | | |
QY 503 QAVPSAGAAGEKKNSQSLWVPOS-----GSPEPESIRNTSRSSSTPAI 547
| : | | |
Db 704 POTSLSAPV-----POSLEPPPALHRSPPDSTEAFDPKRPDPELYEKSr---AI 752
| : | | |
QY 548 FAS---DPATCPIIPGC-----ETIETISKRTGIGLSIVGSD---TLGAFIHEV 594
| : | | |
Db 753 YESRQOVPRTSFRMDSGPDYKELDVLHRMESGFGFRILGDEGQPLIGA-----V 807
| : | | |
QY 595 YEGGAACAKDGRIMAGDQILLEVNGIDLRKATHDEAIVLQTPQ--RVRLTYR----- 645
| : | | |
Db 808 IAMGSADRDGRLHPGDELYVDGIPVAGKTHRYVIDLMHHAARMGQVNLTVRRKYLCEG 867
| : | | |
QY 646 -----DEAPYKE-----EVCDTLIELOKKP 667
| : | | |
Db 868 PCPENGRSPGSVTHHSSPRSDYATYANSNHAAPSNNASPPGCFASHSLQTSVITIRKE 927
| : | | |
QY 668 GKGLGSIYGRN--DTGVFS-----DIVKGIADPDGRLIOGDOILLVNGEDVRNAS 719
| : | | |
Db 928 NGCFGVVIISLNRPESGATITVPHKIGRIIDGSPADRCALKVGRILLAVNGOSIINMP 987
| : | | |
QY 720 QEAVALKCSIGTYLEV-----GRKAGPFHSERRP--SQTQVSEGSLS 764
| : | | |
Db 988 HADIYKLIKDAGLSYTLRIIPEELNPTAPSSEKQSPMAOHSPLAQOHSPLAQPSPA 1047
| : | | |
QY 765 SFTFPLSGSTSESSSKKNALASEIOGLRTVEKKGPTDSLGISIAGVGSPLGDVP 824
| : | | |
Db 1048 TENSPPVAPAPROPQLOLOCHENSYSREVKARQDVK-----PDIRQP 1088
| : | | |
QY 825 IFIAMHPTGVAAGTQKLRVGDRIYITICGSTEGMTHTGAANLKNASGSIEMQVAVAGD 884
| : | | |
Db 1089 PFTDYROP-----PLDYRQPPGCD 1107
| : | | |
QY 885 VGVVT-----GHQEPASSLSFTGLTSTIFQDDLGPPQ---CKSTLERGPDGLGFSI 936
| : | | |
Db 1108 YSQPSPLDYRQHSPTROYPLS-----DYROPQDFDYFTVDMKKGKGFSGFSI 1155
| : | | |
QY 937 VGGYGSPhGDLPIYVKTVPFAKGAASEDRGLKRGDOIIAANGOSLEGVTHEEAVAILKRTK 996
| : | | |
Db 1156 RCG---REYKMDLYVRLAEDGPAINGRVGDQIIEINGESTRDMTHARAIELIKSGG 1212
| : | | |
QY 997 GTVTIMV 1003
| : | | |
Db 1213 RRVRLLL 1219

```

Search completed: July 12, 2001, 14:42:06  
 Job time: 214 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2001, 14:43:17 ; Search time 21.71 Seconds

(without alignments) updates/sec  
1585.756 Million cell

Title: US-09-502-698-2

Perfect score: 5085

Sequence: 1 MLDNWKSEKSEKRTINIAKGN.....EEAVAILIKRTKGVITLWVLS 1005

Scoring table: BLOSUM62

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Database : SWISSPROT\_39:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|-------------|
| 1          | 509   | 10.0        | 2485   | 1  | PTND_HUMAN  |
| 2          | 405.5 | 8.0         | 852    | 1  | DIG2_RAT    |
| 3          | 399   | 7.8         | 817    | 1  | DIG2_HUMAN  |
| 4          | 397.5 | 7.8         | 870    | 1  | DIG2_MOUSE  |
| 5          | 393   | 7.7         | 849    | 1  | DIG3_MOUSE  |
| 6          | 391   | 7.7         | 849    | 1  | DIG3_RAT    |
| 7          | 384   | 7.6         | 767    | 1  | DIG4_HUMAN  |
| 8          | 383   | 7.5         | 724    | 1  | DIG4_MOUSE  |
| 9          | 382   | 7.5         | 724    | 1  | DIG4_RAT    |
| 10         | 379   | 7.5         | 960    | 1  | DIG1_DROME  |
| 11         | 378.5 | 7.4         | 911    | 1  | DIG1_RAT    |
| 12         | 377   | 7.4         | 904    | 1  | DIG1_HUMAN  |
| 13         | 283   | 5.6         | 631    | 1  | IL16_HUMAN  |
| 14         | 243   | 4.8         | 1736   | 1  | ZOI_HUMAN   |
| 15         | 242   | 4.8         | 933    | 1  | ZOI_HUMAN   |
| 16         | 240   | 4.7         | 898    | 1  | ZOI_CANFA   |
| 17         | 237.5 | 4.7         | 1745   | 1  | ZOI_MOUSE   |
| 18         | 235.5 | 4.6         | 905    | 1  | ZOI_MOUSE   |
| 19         | 211   | 4.1         | 1816   | 1  | AF6_HUMAN   |
| 20         | 205.5 | 4.0         | 839    | 1  | APB1_RAT    |
| 21         | 182   | 3.6         | 206    | 1  | OM25_RAT    |
| 22         | 177.5 | 3.5         | 571    | 1  | APB3_MOUSE  |
| 23         | 177   | 3.5         | 837    | 1  | APB1_HUMAN  |
| 24         | 176   | 3.5         | 145    | 1  | OM25_HUMAN  |
| 25         | 173.5 | 3.4         | 569    | 1  | APB3_RAT    |
| 26         | 165   | 3.2         | 982    | 1  | ABPA_CAEEL  |
| 27         | 161.5 | 3.2         | 2842   | 1  | APC_RAT     |
| 28         | 160   | 3.1         | 750    | 1  | APB2_RAT    |
| 29         | 158   | 3.1         | 749    | 1  | APB2_HUMAN  |
| 30         | 157   | 3.1         | 623    | 1  | DSH_DROME   |
| 31         | 153.5 | 3.0         | 2845   | 1  | MAPB_MOUSE  |
| 32         | 153   | 3.0         | 2468   | 1  | MAPB_HUMAN  |
| 33         | 151   | 3.0         | 680    | 1  | APB2_MOUSE  |

## ALIGNMENTS

RESULT 1

| PTND_HUMAN | STANDARD:   | PTT: 2485 AA. | PCGA_CHICK |
|------------|---|---------------|------------|
| AC         | Q12923; Q15263; Q16826; Q15264; Q15265; Q15159;   |               |            |
| AD         | Q12923; Q15263; Q16826; Q15264; Q15265; Q15159;   |               |            |
| DT         | 01-OCT-2000 (Rel. 40, Created)  |               |            |
| DT         | 01-OCT-2000 (Rel. 40, Last sequence update)   |               |            |
| DT         | 01-OCT-2000 (Rel. 40, Last annotation update)   |               |            |
| DE         | PROTEIN TYROSINE PHOSPHATASE, NON-RECEPTOR TYPE 13 (EC 3.1.3.48)  |               |            |
| DE         | (PROTEIN-TYROSINE PHOSPHATASE 1E) (PTP-BAS) (PROTEIN-TYROSINE PHOSPHATASE PTP1) (FAS-ASSOCIATED PROTEIN-TYROSINE PHOSPHATASE 1) (FAP-1).                  |               |            |
| DE         | PTPN13 OR PTP1E OR PTP1 OR PNP1.  |               |            |
| GN         | Homo sapiens (Human).   |               |            |
| OS         | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.                               |               |            |
| OC         | NCBI_TaxID=9606;  |               |            |
| OX         | [1]   |               |            |
| RN         | SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.   |               |            |
| RP         | TISSUE=Leukemia;  |               |            |
| RC         | TISSUE=Breast carcinoma;  |               |            |
| RC         | MEDLINE=94350988; PubMed=8071359;   |               |            |
| RA         | Banville D., Ahmed S., Stocco R., Shen S.-H.;   |               |            |
| RT         | "A novel protein-tyrosine phosphatase with homology to both the cytoskeletal proteins of the band 4.1 family and junction-associated guanylate kinases."; |               |            |
| RT         | J. Biol. Chem. 269:22320-22327(1994).   |               |            |
| RT         | [2]   |               |            |
| RN         | SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.   |               |            |
| RP         | TISSUE=Leukemia;  |               |            |
| RC         | TISSUE=Breast carcinoma;  |               |            |
| RC         | MEDLINE=94116679; PubMed=8287977;   |               |            |
| RA         | Maekawa K., Imagawa N., Nagamatsu M., Harada S.;  |               |            |
| RT         | "Molecular cloning of a novel protein-tyrosine phosphatase containing a membrane-binding domain and GGF repeats.";  |               |            |
| RT         | FEBS Lett. 337:200-206(1994).   |               |            |
| RT         | [3]   |               |            |
| RN         | SEQUENCE FROM N.A.  |               |            |
| RP         | TISSUE=Fibroblast;  |               |            |
| RC         | MEDLINE=95014139; PubMed=7929060;   |               |            |
| RA         | Sares J., Claesson-Welsh L., Heldin C.-H., Gonen L.J.;  |               |            |
| RT         | "Cloning and characterization of PTP1, a protein tyrosine phosphatase with similarities to cytoskeletal-associated proteins.";                            |               |            |
| RT         | J. Biol. Chem. 269:24082-24089(1994).   |               |            |
| RT         | [4]   |               |            |
| RN         | SEQUENCE OF 1216-2490 FROM N.A.   |               |            |
| RP         | TISSUE=Pancreas;  |               |            |
| RC         | Wang H.Y.;  |               |            |
| RA         | Submitted (JUN-1994) to the EMBL/GenBank/DBJ databases.   |               |            |
| RN         | [5]   |               |            |
| RP         | STRUCTURE BY NMR OF 1361-1456.  |               |            |
| RC         | MEDLINE=20170882; PubMed=10704206;  |               |            |
| RA         | Kozlov G., Gehring K., Ekkel I.;  |               |            |
| RT         | "Solution structure of the PDZ domain from human phosphatase hPTP1E and its interactions with C-terminal peptides from the Fas receptor.";                |               |            |
| RT         | Biochemistry 39:2572-2580(2000).  |               |            |
| CC         | -I- FUNCTION: BINDS TO A NEGATIVE REGULATORY DOMAIN IN FAS THAT INHIBITS FAS-INDUCED APOPTOSIS.   |               |            |

CC -1- CATALYTIC ACTIVITY: PROTEIN TYROSINE PHOSPHATE + H(2)O - PROTEIN  
 CC TYROSINE + PHOSPHATE.  
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).  
 CC -1- ALTERNATIVE PRODUCTS: 3 ISOFORMS: 1 (SHOWN HERE), 2 AND 3; ARE  
 CC PRODUCED BY ALTERNATIVE SPLICING.  
 CC -1- TISSUE SPECIFICITY: PRESENT IN MOST TISSUES WITH THE EXCEPTION OF  
 CC THE LIVER AND SKELETAL MUSCLE. MOST ABUNDANT IN LUNG, KIDNEY AND  
 CC FETAL BRAIN.  
 CC -1- SIMILARITY: CONTAINS A DOMAIN FOUND IN BAND 4.1, EZRIN, MOESIN,  
 CC RADIAXIN, AND TALIN.  
 CC -1- SIMILARITY: CONTAINS 5 PDZ/DHR DOMAINS.  
 CC -1- SIMILARITY: BELONGS TO THE NON-RECEPTOR CLASS OF THE PROTEIN-  
 CC TYROSINE PHOSPHATASE FAMILY.  
 CC -----  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: U12128; AAB60339.1; -  
 DR EMBL: D21209; BAA04750.1; -  
 DR EMBL: D21210; BAA04751.1; -  
 DR EMBL: D21211; BAA04752.1; -  
 DR EMBL: X80289; CAA56563.1; -  
 DR EMBL: X79676; CAA56124.1; -  
 DR PDB: 3PDZ; 17-MAR-00.  
 DR MIM: 600267; -  
 DR Interpro: IPR000242; -  
 DR Interpro: IPR000299; -  
 DR Interpro: IPR000387; -  
 DR Interpro: IPR001478; -  
 DR Pfam: PF00102; Y-phosphatase; 1.  
 DR Pfam: PF00373; Band\_41.1.  
 DR Pfam: PF00393; PDZ; 5.  
 DR PRINTS: PR00700; PRTYPHPTASE.  
 DR PRINTS: PR00935; BANDA1.  
 DR PROSITE: PS00660; BAND\_41.1; FALSE\_NEG.  
 DR PROSITE: PS00661; BAND\_41.2; FALSE\_NEG.  
 DR PROSITE: PS50057; BAND\_41\_3; 1.  
 DR PROSITE: PS50106; PDZ; 5.  
 DR PROSITE: PS50383; TYR\_PHOSPHATASE\_1; FALSE\_NEG.  
 DR PROSITE: PS50056; TYR\_PHOSPHATASE\_2; 1.  
 DR PROSITE: PS50055; TYR\_PHOSPHATASE\_PP; 1.  
 DR Structural protein: Cytoskeleton; Hydrolase; Repeat; 3D-structure;  
 KM Alternative splicing; Coiled coil.  
 FT DOMAIN 36  
 FT DOMAIN 59  
 FT DOMAIN 585  
 FT DOMAIN 879  
 FT DOMAIN 2237  
 FT DOMAIN 2485  
 FT DOMAIN 379  
 FT DOMAIN 399  
 FT DOMAIN 469  
 FT DOMAIN 504  
 FT DOMAIN 1775  
 FT DOMAIN 1804  
 FT DOMAIN 2057  
 FT DOMAIN 2085  
 FT DOMAIN 1093  
 FT DOMAIN 1178  
 FT DOMAIN 1368  
 FT DOMAIN 1452  
 FT DOMAIN 1501  
 FT DOMAIN 1588  
 FT DOMAIN 1788  
 FT DOMAIN 1868  
 FT DOMAIN 1882  
 FT DOMAIN 1965  
 FT DOMAIN 1742  
 FT DOMAIN 1749  
 FT ACT\_SITE 2408  
 FT ACT\_SITE 2408  
 FT VARSPIC 884  
 FT VARSPIC 1074  
 FT VARSPIC 1056  
 FT VARSPIC 1074  
 FT CONFLICT 1134  
 FT CONFLICT 1135  
 FT CONFLICT 1216  
 FT CONFLICT 1229  
 FT CONFLICT 1238  
 FT CONFLICT 1239  
 FT CONFLICT 1357  
 FT CONFLICT 1357  
 FT CONFLICT 1362  
 FT CONFLICT 1363  
 FT CONFLICT 1383  
 FT CONFLICT 1383  
 FT CONFLICT 1538  
 FT CONFLICT 1538  
 FT CONFLICT 1649  
 FT CONFLICT 1649

FT CONFLICT 1698 1714 KSOEDTCTMEYYPQKI -> RVKKIPEVCFILTKRK  
 FT CONFLICT 1797 1797 G -> A (IN REF. 4).  
 FT CONFLICT 1856 1857 AA -> G (IN REF. 4).  
 FT CONFLICT 2069 2069 A -> S (IN REF. 4).  
 FT CONFLICT 2206 2210 GLDQ -> VARS (IN REF. 4).  
 SQ SEQUENCE 2485 AA; 276903 MM; 801B31597C66962B CRC64;  
 Query Match 10.08; Score 509; DB 1; Length 2485;  
 Best Local Similarity 22.88; Pred. No. 1,1e-19;  
 Matches 243; Conservative 140; Mismatches 371; Indels 314; Gaps 37;  
 QY 162 EPSKSLGISYGRGMSRLNCEVMMRGIFIKHYLESPACKNGTLKPGDRIVAPSQE 221  
 DB 990 EPPQQTVAELVGRKSHOMSRDAESLAGV---TLNNKSVASLNRPERKHSDDSSST 1046  
 QY 222 SEPEKAPLCVPPPPPSAFAMEGSDHTQSSASKISODVDKDEFGYSWKIRERYGTLTG 281  
 DB 1047 EDPQAVYLGWT-----MHSSGNSSSQVPLKENDV-----LHKRMYSVSS 1086  
 QY 282 ---ELHMLELK-GHSGLGISLAKNDRSR--MSVFIYGDIPNCAAGKDRLOJADELLE 335  
 DB 1087 PEREITLVNLRKAKKAYGLGEOIIGEEKMGRDLGIFTSVAPGPPADLDGLKPGRLIS 1146  
 QY 336 INCOILYGRSHONASIIKCAPSKYKIIIFIRNKDAVQMAVCPGNAVEPLPSNSENQNK 395  
 DB 1147 VNSVSLGVSHPHAAIELONAPEDVTLVISOPKRIKIKVPSTPYH-----LTNEKNKNTMK 1202  
 QY 396 ETEPTVTTSDAANDLSFEKN-----VOHLELPK--DQGL--GIAISEDTLSGVYIKS 445  
 DB 1203 SS---YMODSAIDSSSKDHMSRGLRHISENSFGSPGSLREGSLSSODSRRESASLSQ 1258  
 QY 446 LTEHGVATOGRLKVGQOILAVDEIVVGYPIEFISLTKAKTVTLTHAENPQAV 505  
 DB 1259 SQVNGFFAS---HLGQITW---QESQHGSPSPVIS-----KATEKETITDSNQSTKK 1306  
 QY 506 PSAAGASGEKK-----NSQSILMPQSGSPESPESIRNTRSSPTAIFASDPATC 555  
 DB 1307 PGIDVDVDSRGSDMDDEATYSSSDOHQPKQ---ESSSVNNTSNKMNKTFSSSPK- 1362  
 QY 556 PIRGCEETITISKGRGLGISYVGSDDL--GAFIIEHYEGCAKCGRLMAGQIL 613  
 DB 1363 ---PGDIFEVELAKNDNSLISYGVGVNTSVRHGIVYKAIIPGAAESDGRIRIKGDRVL 1419  
 QY 614 EVNGIDIRKATHDAIVNLTROPVRRLTYRDAAPYKEEV---C----- 656  
 DB 1420 AVNGVSLGATHKQAVETLRNTGQVHLLLEKQSPISKEHVPYTPQCTTSDQNAQGP 1479  
 QY 657 -----DTLTIELQKPGKGLS-----IVGRNDYGVFVSDIVK 691  
 DB 1480 EKVKKTQVKQYSVTEENFEVLFKN--SGGLGFSRREDNLPEQINASTIVKVLFP 1538  
 QY 692 GGIADPGRLIQGQIILVNGEDVRNASEVNAALCISLCTVLEGRKAG----- 744  
 DB 1539 GQPAAESGKIDVDVILKVGASLKLGSQOEVISALRGTAPEVLLCRPPGVLPEIDT 1598  
 QY 745 ---PRHSE-----RRPSQVSGES-- 762  
 DB 1599 ALLTPLOSPAOVLPNSSKDSQSPCEOSTSSDENEMSKKCKSPSRDSTSDSSGS 1658  
 QY 763 -----LSSFP-----PLGSGST----- 775  
 DB 1659 GEDDOLVAPANISMSTWSALHQTLSMNVSOAQSHHAPASQEDTICTMYYPPKIPNK 1718  
 QY 776 -----SESLSSSK-----KNALASEI 792  
 DB 1719 EFEDSNPSPLPPDMAPOQSOSESASSSMQYHIHHSIPTRODNWPLKLDLNLH 1778  
 QY 793 OGRL-----IYEMKKKPTDSIGISAG---VSGPLDDVPIFIAMHPTGVAQOTOKLRY 844  
 DB 1779 EDELEVELLITLIKSEKSGSLGFTVYKGNORIGCYHADV-----IQDP---ANSQRLKP 1830





QY 997 GTVTLM 1002  
 DB 492 QVTIIL 497

RESULT 3  
 ID DLG3\_HUMAN STANDARD; PRT; 817 AA.

AC 092796; 09UL18;  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 01-OCT-2000 (Rel. 40, Last annotation update)  
 DE PRENAFTIC PROTEIN SAPIOT (SYNAPSE-ASSOCIATED PROTEIN 102)  
 GN DE (NEUROENDOCRINE-DLG) (NE-DLG) (DISCS, LARGE HOMOLOG 3).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
 NX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Fetal brain;  
 RX MEDLINE=97332623; Pubmed=9188857;  
 RA Makino K., Kuwahara H., Masuko N., Nishiyama Y., Morisaki T.,  
 RA Sasaki J., Nakao M., Kuwano A., Nakata M., Ushio Y., Saya H.,  
 RT Cloning and characterization of NE-dlg: a novel human homolog of the  
 RT Drosophila discs large (dlg) tumor suppressor protein interacts with  
 RT the APC protein.;  
 RL Oncogene 14:2425-2433(1997).  
 RN [2]  
 RP SEQUENCE OF 330-817 FROM N.A.  
 RC TISSUE=Brain;  
 RX MEDLINE=20039619; Pubmed=10574462;  
 RA Nagase T., Ishikawa K.-I., Kikuno R., Hirosewa M., Nomura N.,  
 RA Ohara O.;  
 RT Prediction of the coding sequences of unidentified human genes. XV.  
 RT The complete sequences of 100 new cDNA clones from brain which code  
 RT for large proteins in vitro.;  
 RL DNA Res. 6:337-345(1999).  
 CC - FUNCTION: INTERACTS WITH THE CYTOPLASMIC TAIL OF THE NMDA RECEPTOR  
 CC SUBUNIT NR2B (BY SIMILARITY).  
 CC - SIMILARITY: CONTAINS 1 SH3 DOMAIN.  
 CC - SIMILARITY: CONTAINS 3 PDZ/DHR DOMAINS.  
 CC - SIMILARITY: BELONGS TO THE MAGUK FAMILY OF CELL JUNCTION PROTEINS.  
 CC - SIMILARITY: BELONGS TO THE MAGUK FAMILY OF CELL JUNCTION PROTEINS.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: U49089; AAB61453.1; -  
 CC EMBL: AB033058; BAB86546.1; -  
 CC HSSP: Q12959; 1PDR.  
 CC MIM: 300189; -  
 CC InterPro: IPR000619; -  
 CC InterPro: IPR001452; -  
 CC InterPro: IPR001478; -  
 CC Pfam: PF00625; Guanylate\_kin. 1;  
 CC Pfam: PF00595; PDZ; 3.  
 CC Pfam: PF00018; SH3; 1.  
 CC PROSITE: PS00856; GUANYLATE\_KINASE\_1; 1.  
 CC PROSITE: PS50052; GUANYLATE\_KINASE\_2; 1.  
 CC PROSITE: PS50106; PDZ; 3.  
 CC PROSITE: PS50002; SH3; 1.  
 CC SH3 domain; Repeat.  
 CC FT DOMAIN 130 217 PDZ 1.  
 CC FT DOMAIN 226 311 PDZ 2.  
 CC FT DOMAIN 379 465 PDZ 3.  
 CC FT DOMAIN 503 568 SH3.

FT DOMAIN 628 803 GUANYLATE KINASE.  
 FT CONFLICT 330 381 FTALADNHISNSLIGYLGAVESKVSYPAPQVPTPTSP  
 FT CONFLICT 330 381 PRHMLAEEDFT -> AARREGAMERARKSGSLAMGLS  
 FT CONFLICT 592 606 ASASARRASQRMAMPRLSRPGDA (IN REF. 2).  
 FT CONFLICT 592 606 DFGSLDDYIGAKNL -> SIKTKRKSPFLSKPFPYASK  
 FT CONFLICT 592 606 ENNAOESSIOEQVTSNTSDSESS (IN REF. 2).  
 SO SEQUENCE 817 AA; 90344 MW; 3D7512EC4713FC4E CRC64;

Query Match 7.88; Score 399; DB 1; Length 817;  
 Best Local Similarity 24.98; Pred. No. 2.1e-14;  
 Matches 132; Conservative 77; Mismatches 168; Indels 154; Gaps 16;

QY 502 SQAVPSAAGSAGEKKNSSOSLWPOGSPESIRTSSTSPAIAPSPATPIPGC 561  
 DB 56 SQTLPQAGATPPRR---TKAKLIPTRGDVGPVLKVPVCKSPKLGSGSPMW---PEC 109  
 QY 562 ETT-----IEISKRTGLSLYGSDDL---LCAFIHEYEE 597  
 DB 110 TCTNRDWEQVNGSDGMFKYEIVLERGNSGLGFSIAGIDNPVDDPCIFIT-KIIPG 168  
 QY 598 GAACKDGRMAAGDILEVNGIDLRKATHDEAINVLRQTPQVRLLTRDEAPYKEEVED 657  
 DB 169 GAAMDRLGVNDVCLVKNVEVEVSHSRVLEAKGAPVRLVRRRQPPPE---T 223  
 QY 658 TLTELOKKRKGKGLSLYVK-----RNDGVFSDIVKGIADPPGRILQDQILLVN 711  
 DB 224 IMEVNLLKGP-KGLGFSIAGIGNQHTIPGDSIYITITIEGGAQKGRIGORLLAVN 282  
 QY 712 GEDVRNASEVVAALLCSIGTYLVEGRYKAGFPFHEBRPSQTSQVSEGLSFTPLS 771  
 DB 283 NTNLQDVHEEVAASLNTSDMYLKVA--KPSGLH-----LNDMYVAPPD 325  
 QY 772 GSSTSESLSSSKKNALASEIQLRTVEKKGPTDSIGISIAGVSGPLDVPFIAMMH 831  
 DB 326 YASTFTLADNH----- 337  
 QY 832 PTGVAAQTKLRVGRDRIYITIGISTEGMTHTQAVNLLKNSGSIEMOVVAGDVSVTGH 891  
 DB 338 -----ISNSSLGYL-----GAESKRV-----SY 356  
 QY 892 HQEPASSLSPTGLTSTISINODLGPPQCKSITLERPDGSGFSGVSPHGLPIYV 951  
 DB 357 PAPQVPTPTSPTRISLPRMLAEEDF-TREPKTILHKSTGEGFIVGG---EDEE-GIFV 411

QY 952 KTVFAKGAASEDEGRKRGDQIIAANGSLGVTHEEVAALLKRTKGTVTLM 1002  
 DB 412 SFLIAGPADLSGELRGRDRLTSVNGVNLNATHEGAAALAKRAGOSVTIV 462

RESULT 4  
 ID DLG2\_HUMAN STANDARD; PRT; 870 AA.

AC 015700;  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 01-OCT-2000 (Rel. 40, Last annotation update)  
 DE CHANNEL ASSOCIATED PROTEIN OF SYNAPSE-110 (CHAPSIN-110) (DISCS, LARGE  
 GN HOMOLOG 2).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
 NX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC MEDLINE=96310881; Pubmed=8755482;  
 RA Kim E., Cho K.-O., Rothschild A., Sheng M.;  
 RT Heteromultimerization and NMDA receptor-clustering activity of  
 RT Chapsyn-110, a member of the PSD-95 family of proteins.;  
 RL Neuron 17:103-113(1996).  
 CC - FUNCTION: INTERACTS WITH THE CYTOPLASMIC TAIL OF NMDA RECEPTOR  
 CC SUBUNITS AS WELL AS POTASSIUM CHANNELS.

| Query Match               | Score | DB 1              | Length              |
|---------------------------|-------|-------------------|---------------------|
| Best Local Similarly      | 7.88  | 397.5             | 870                 |
| Matches 154; Conservative | 28.18 | Pred. No. 2.8e-14 |                     |
|                           | 69    | Mismatches 205    | Indels 121; Gaps 23 |

|    |     |           |                      |                   |                |     |
|----|-----|-----------|----------------------|-------------------|----------------|-----|
| Db | 433 | FNIVGG--- | EDGE-GIFVPSFILAGPADL | SGSELRGQDILSYNGID | LRGASHGEQAAALK | 488 |
| Qy | 994 | RTKGTVTLM | 1002                 |                   |                |     |
|    |     | :         |                      |                   |                |     |
| Db | 489 | GAGQTVTII | 497                  |                   |                |     |

|    |                                       |   |
|----|---------------------------------------|---|
| CC |                                       |   |
| DR | EMBL; D87117; BAAL3249.1;             | -.                                      |
| DR | HSSP; O12959; IPDR.                   |   |
| DR | MCD; MGI:1888986; DlgH3.              |   |
| DR | InterPro; IPR000619;                  | -.                                      |
| DR | InterPro; IPR001452;                  | -.                                      |
| DR | InterPro; IPR001478;                  | -.                                      |
| DR | Pfam; PF00625; Guanylate_kin;         | 1.                                      |
| DR | Pfam; PF00595; PDZ_3.                 |   |
| DR | Pfam; PF00018; SH3_1.                 |   |
| DR | PROSITE; PS00856; GUANYLATE_KINASE_1; | 1.                                      |
| DR | PROSITE; PS50052; GUANYLATE_KINASE_2; | 1.                                      |
| DR | PROSITE; PS50106; PDZ_3.              |   |
| DR | PROSITE; PS50002; SH3_1.              |   |
| KW | SH3 domain; Repeat.                   |   |
| FT | DOMAIN                                | 149 .. 235 PDZ 1.                       |
| FT | DOMAIN                                | 244 .. 330 PDZ 2.                       |
| FT | DOMAIN                                | 404 .. 484 PDZ 3.                       |
| FT | DOMAIN                                | 519 .. 589 SH3.                         |
| FT | DOMAIN                                | 659 .. 849 GUANYLATE KINASE.            |
| SQ | SEQUENCE                              | 849 AA; 93482 MW; EF3EFD7I5338EE CRC64; |

[illegible]



RESULT 7  
 DLG4\_HUMAN STANDARD: PRT: 767 AA.  
 AC P78352; Q92941; Q9URK8;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 30-NOV-2000 (Rel. 39, Last sequence update)  
 DT 01-OCT-2000 (Rel. 40, Last annotation update)  
 DE PRESYNAPTIC DENSITY PROTEIN 95 (PSD-95) (DISCS, LARGE HOMOLOG 4).  
 GN DLG4 OR PSD95.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_Taxid=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-Mammary gland;  
 RA MEDLINE=97432822; PubMed=9286702;  
 RA Stathakis D.G., Hoover K.H., You Z., Bryant P.J.;  
 RT "Human postsynaptic density-95 (PSD95): location of the gene (DLG4)  
 RT and possible function in nonneural as well as in neural tissues.";  
 RL Genomics 44:71-82(1997).  
 RN [2]  
 RP REVISIONS.  
 RC TISSUE-Mammary gland;  
 RA Stathakis D.G., Hoover K.H., You Z., Bryant P.J.;  
 RL Submitted (JUL-1998) to the EMBL/Genbank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA Stathakis D.G., Udar N., Sandgren O., Andreasson S., Bryant P.J.,  
 RA Small K., Forzani-Semb K.;  
 RT "Genomic organization of human DLG4, the gene encoding post-synaptic  
 RT density 95 (PSD95).";  
 RL J. Neurochem. 73:0-0(1999).  
 RN [4]  
 RP SEQUENCE OF 81-401 FROM N.A.  
 RC TISSUE-Brain;  
 RA Breman J.E., Brecht D.S., Parkinson J.F., Manzana W.P., McClary J.A.;  
 RL Submitted (AUG-1996) to the EMBL/Genbank/DBJ databases.  
 CC -1- FUNCTION: INTERACTS WITH THE CYTOPLASMIC TAIL OF NMDA RECEPTOR  
 CC SUBUNITS. MAY BE INVOLVED IN SYNAPTIC PLASTICITY.  
 CC -1- SUBCELLULAR LOCATION: CONCENTRATED AT SYNAPTIC JUNCTIONS PRIMARILY  
 CC ON THE PRESYNAPTIC SIDE (WAS ORIGINALLY THOUGHT TO BE  
 CC POSTSYNAPTIC).  
 CC -1- TISSUE SPECIFICITY: PRESYNAPTIC DENSITY FRACTION OF BRAIN.  
 CC -1- SIMILARITY: CONTAINS 3 PDZ/DRH DOMAINS.  
 CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.  
 CC -1- SIMILARITY: CONTAINS A GUANYLATE KINASE-LIKE DOMAIN.  
 CC -1- SIMILARITY: BELONGS TO THE MAGUK FAMILY OF CELL JUNCTION PROTEINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 DR EMBL: U83192; AAC52113.1;  
 DR EMBL: AF156495; AAD56173.1;  
 DR EMBL: U68138; AAB07736.1;  
 DR HSSP: Q12959; 1PDR.  
 DR MIM: 602887;  
 DR InterPro: IPR000619;  
 DR InterPro: IPR001452;  
 DR InterPro: IPR001478;  
 DR Pfam: PF00625; Guanylate\_kin; 1.  
 DR Pfam: PF00395; PDZ; 3.  
 DR Pfam: PF00018; SH3; 1.  
 DR PROSITE: PS00856; GUANYLATE\_KINASE\_1; 1.  
 DR PROSITE: PS50052; GUANYLATE\_KINASE\_2; 1.  
 DR PROSITE: PS50106; PDZ; 3.  
 DR PROSITE: PS50002; SH3; 1.  
 KW SH3 domain; Repeat.

FT DOMAIN 108 194 PDZ 1.  
 FT DOMAIN 203 289 PDZ 2.  
 FT DOMAIN 356 436 PDZ 3.  
 FT DOMAIN 471 541 SH3.  
 FT DOMAIN 577 767 GUANYLATE KINASE.  
 FT CONFLICT 46 46 E -> V (IN REF. 3).  
 FT CONFLICT 81 83 VIV -> EFR (IN REF. 4).  
 FT CONFLICT 399 401 GDO -> AGI (IN REF. 4).  
 SQ SEQUENCE 767 AA; 85429 MW; BE1019159E65B2D8 CRC64;  
 Query Match 7.6%; Score 384; DB 1; Length 767;  
 Best Local Similarity 25.8%; Pred. No. 1.2e-13;  
 Matches 131; Conservative 76; Mismatches 153; Indels 148; Gaps 18;  
 QY 518 NSSQSLMVPQSGSP---EPESIRNTSRSSPTAIFASDPATPC---IIPGCTTIE--- 566  
 DB 50 SESQKRYQDEDTPTPLEHSPALHPNOA-NSPVIVNTDTLEAPGYELGVNGTEGEYE 108  
 QY 567 --ISKRGNGIGSTIVGSGPT-LLG---AFIHEVEEGACAKDGRUMGDOILEVNGIDL 620  
 DB 109 ITLERKNSGLFSLGIAGGTDNPHIGDPSFTFKIIPGCAAADGRLRVNDSTLFPNEVDV 168  
 QY 621 KRATDEAINVLRQTPORVRLTLVYDEAPYKEEYCDTLTIELOKPKGGLSTIVGK-- 678  
 DB 169 REVTHSAANVEALKEAGSTVRLVYMRKPPAEK-----VMEIKLIGP-KGLGFSIAGGVG 222  
 QY 679 ---RNDTGVESVDIVYKGIADPGRLIOGDOIILVNGEDVRNASOEAVALLKCSLGTIV 734  
 DB 223 NOHIPGDNSTIVYTKIIEGGAARDKRLQIGDKILAVNSVGLDEVDHVEDVAALKMTYDVV 282  
 QY 735 TLEVGRIKAGPPHSEKRPQTSQVEGSLSTFPLSGSTSESSSKNAALSEIOLG 794  
 DB 283 YLKVA-----KPS-NAVLSQ---STAPDITTSYSDHLDNEISHSS----- 319  
 QY 795 LRTVEKKKPTDLSIGISINGVSPGLGVPIFIAMHPGVAAQOTKLKLVGDRIYTCGT 854  
 DB 320 -----YLGNDYPT----- 327  
 QY 855 STEGMTHIOAVLLKNASSIEMQVAGDSVVTGHHOEPASSLSLSTGLTSTIPDD 914  
 DB 328 ---AMPT-----SPRRSPVAKDLGIED 349  
 QY 915 LGPPCKSTTLRGPDPGIGSTIVGSGPHGDLPIYKTVFAKGAASEGRLKRGQIIA 974  
 DB 350 I-PREERRIVIRGSTGLFNIVG---EDGE-GIFISFLIAGGPDLSGELKKGQIIIS 404  
 QY 975 VNGOSLEGVTHEEAVAILKRTKGTVTLM 1002  
 DB 405 VNGVDLRNASHEDQALALKNAQGTVII 432  
 RESULT 8  
 DLG4\_RAT STANDARD: PRT: 724 AA.  
 AC P31016; P97631;  
 DT 01-JUL-1993 (Rel. 26, Created)  
 DT 01-JUL-1993 (Rel. 26, Last sequence update)  
 DT 01-OCT-2000 (Rel. 40, Last annotation update)  
 DE PRESYNAPTIC DENSITY PROTEIN 95 (PSD-95) (PRESYNAPTIC PROTEIN SAP90)  
 DE (SYNAPTIC-ASSOCIATED PROTEIN 90) (DISCS, LARGE HOMOLOG 4).  
 GN DLG4 OR PSD95.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_Taxid=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=SPRAGUE-DAWLEY; TISSUE=Brain;  
 RA MEDLINE=93040233; PubMed=1419001;  
 RA Cho K.-O., Hunt C.A., Kennedy M.B.;  
 RT "The rat brain postsynaptic density fraction contains a homolog of  
 RT the Drosophila discs-large tumor suppressor protein.";

```

RL Neuron 9:929-942(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-SPRAGUE-DAWLEY; TISSUE=Brain;
RX MEDLINE=93186749; PubMed=7680343;
RA Kistner U., Wenzel B.M., Voh R.W., Cases-Langhoff C., Garner A.M.,
RT Appelbauer U., Voss B., Gundelfinger E.D., Garner C.C.;
RT "SAP90, a rat presynaptic protein related to the product of the
RT Drosophila tumor suppressor gene dlg-A.";
RL J. Biol. Chem. 268:4580-4583(1993).
RN [3]
RP SEQUENCE OF 566-625 FROM N.A.
RC STRAIN=Wistar Kyoto; TISSUE=vascular smooth muscle;
RA Adams L.D., Werny I., Schwartz S.M.;
RL Submitted (NOV-1996) to the EMBL/Genbank/DBJ databases.
RN [4]
RP X-RAY CRYSTALLOGRAPHY (1.82 ANGSTROMS) OF 302-402.
RX MEDLINE=96270509; PubMed=8674113;
RA Doyle D.A., Lee A., Lewis J., Kim E., Sheng M., Mackinnon R.;
RT "Crystal structures of a complexed and peptide-free membrane protein-
RT binding domain: molecular basis of peptide recognition by PDZ".
RL Cell 85:1067-1076(1996)
CC -1- FUNCTION: INTERACTS WITH THE CYTOPLASMIC TAIL OF NMDA RECEPTOR
CC SUBUNITS. MAY BE INVOLVED IN SYNAPTOGENESIS.
CC -1- SUBCELLULAR LOCATION: CONCENTRATED AT SYNAPTIC JUNCTIONS PRIMARILY
CC ON THE PRESYNAPTIC SIDE (WAS ORIGINALLY THOUGHT TO BE
CC POSTSYNAPTIC).
CC -1- TISSUE SPECIFICITY: PRESYNAPTIC DENSITY FRACTION OF BRAIN.
CC -1- SIMILARITY: CONTAINS 3 PDZ/DRH DOMAINS.
CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
CC -1- SIMILARITY: CONTAINS A GUANYLATE KINASE-LIKE DOMAIN.
CC -1- SIMILARITY: BELONGS TO THE MAGUK FAMILY OF CELL JUNCTION PROTEINS.
CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M96853; AAA41971.1; -.
DR EMBL; X66474; CAA47103.1; -.
DR EMBL; U77090; AAB38270.1; -.
DR PIR; S26407; S26407.
DR PIR; JH0800; JH0800.
DR PDB; 1BE9; 2I-OCT-98.
DR PDB; 1BF6; 2I-OCT-98.
DR InterPro; IPR000619; -.
DR InterPro; IPR001452; -.
DR InterPro; IPR001478; -.
DR Pfam; PF00625; Guanylate_kin; 1.
DR Pfam; PF00595; PDZ; 3.
DR Pfam; PF00018; SH3; 1.
DR PROSITE; PS00856; GUANYLATE_KINASE_1; 1.
DR PROSITE; PS50052; GUANYLATE_KINASE_2; 1.
DR PROSITE; PS50106; PDZ; 3.
DR PROSITE; PS50002; SH3; 1.
KW SH3 domain; Repeat; 3D-structure.
FT DOMAIN 65 151 PDZ 1.
FT DOMAIN 160 246 PDZ 2.
FT DOMAIN 313 393 PDZ 3.
FT DOMAIN 428 498 SH3.
FT DOMAIN 534 724 GUANYLATE KINASE.
FT CONFLICT 61 61 M->L (IN REF. 2).
FT CONFLICT 78 78 S->T (IN REF. 2).
FT CONFLICT 177 182 GVGNOH->ALGTSI (IN REF. 2).
FT CONFLICT 200 200 A->G (IN REF. 2).
FT CONFLICT 254 254 S->T (IN REF. 2).
FT CONFLICT 539 555 ILGPTKQANDDLSEF->ISLDPKTPTVMTIFSSS
FT (IN REF. 2).
FT CONFLICT 623 625 GKH->RDO (IN REF. 3).
FT SEQUENCE 724 AA: 80465 MW: 792204E8E0F9AD85 CCK64;

```

```

Query Match similarity      7.5%: Score 383; DB 1; Length 724;
Best Local Similarity      26.2%: Pred No.1,3e-13;
Matches 128; Conservative  74; Mismatches 143; Indels 144; Gaps 17.

QY 533 PESIRNTSRSSPTAIFASDPATCP-----IIPCEETIE-----ISKRTGLSLSTVGSD 583
      ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 26 PAHLPNOA--NSPPVIVNTDTLEAPGELQVNGTEGEMEEYITLERGNGSLFGSLINGD 84
      ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 584 T-LLG---AFITHEVEEBAACKDRGLYAGDQILLFVNGIDLKKAHDEALNVLROTPQV 639
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 85 NPHIGDPSIFETKTIIPGGAADQGLRVNDSILFVNEVDVREYTHSAVAELKKEGSI 144
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 640 RLTLRYDEAPRYKEEEVCDLTLELQKKPKGGLSLIVGK-----RNDGVFVSDIVKGG 693
      ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 145 RLTYMRKRPRPK-----VMEIKLTKGP-KGLGFSIAGVGNGHITPGDNIYVTAKIEEG 198
      ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 694 IADPGRLIQDQILLVNGEDVYRNASQEAVALLCSLCTVTELEWGRIKAGPFHSERRPS 753
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 199 AAHKRGRLDIGKRIILAVNSGLEVDVHEDAVALAKNTKYDVYLKVA-----RPS 247
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 754 QRSQVSEGLSEFTPLSSSSSTSESLSSSKNALASELQGLRYTEMKKGPTDSIGISA 813
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 248 -NAVYSD-----SYAPDITTSYQHLNDNISHS----- 276
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 814 GGVGSPGLDVPPIFIAMHPTGVAAOTOKRLVRDRIYITIGTSTEGMTHTOAVNLKMASG 873
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 277 -----YLQTDVPT-----AMTP----- 289
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 874 SIEMQVAGGVSVYTGHHQEPASSLSFTGLTSTISFODDGLPPQCKSTLERPDGLG 933
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 290 -----SPRRASIPAKDLGEEDL-PEPRPRIVYHRSSTGG 324
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 934 FSIVGSGSPHGDLPIYKYVTFPAKGAASEDRGLRKGDQIIAVNGSLSEGVTHEEVAALIK 993
      ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : ||
DB 325 FNIWVG---EDGE-GIFISFIAGSPADLSGLRGRDQILSNVGDVLRNASHQAAIALIK 380
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 994 RTKGTVTLMA 1002
      ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 381 NAGQVTVTII 389

RESULT 9
DLG4_MOUSE
AC 062108; STANDARD: PRT: 724 AA.
DB 01-NOV-1997 (Rel. 35, Created)
DB 01-NOV-1997 (Rel. 35, Last sequence update)
DB 01-OCT-2000 (Rel. 40, Last annotation update)
DE PRESYNAPTIC DENSITY PROTEIN 95 (PSD-95) (PRESYNAPTIC PROTEIN SAP90)
DE (SYNAPSE-ASSOCIATED PROTEIN 90) (DISCS, LARGE HOMOLOG 4).
GN DLG4 OR DLG4 OR PSD95.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RX NCBI_TaxID=10090; [1]
RN SEQUENCE FROM N.A.
RC STRAIN=DBA/2; TISSUE=Brain;
RA Kojima N., Yagi T.;
RL Submitted (May-1995) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: INTERACTS WITH THE CYTOPLASMIC TAIL OF NMDA RECEPTOR
CC SUBUNTS. MAY BE INVOLVED IN SYNAPTOGENESIS (BY SIMILARTY).
CC -1- SUBCELLULAR LOCATION: CONCENTRATED AT SYNAPTIC JUNCTIONS PRIMARILY
CC ON THE PRESYNAPTIC SIDE (BY SIMILARTY).
CC -1- SIMILARITY: CONTAINS 3 PDZ/DHR DOMAINS.
CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
CC -1- SIMILARITY: BELONGS TO THE MAGUK FAMILY OF CELL JUNCTION PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

```

the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: D50621; BAA09297.1; -  
DR HSSP: Q12959; 1PDR.  
DR MGP: MGI:1277959; D1gh4.  
DR InterPro: IPR000619; -  
DR InterPro: IPR001452; -  
DR Pfam: PF00625; Guanylate\_kin; 1.  
DR Pfam: PF00595; PDZ; 3.  
DR Pfam: PF00018; SH3; 1.  
DR PROSITE: PS00856; GUANYLATE\_KINASE\_1; 1.  
DR PROSITE: PS50052; GUANYLATE\_KINASE\_2; 1.  
DR PROSITE: PS50106; PDZ; 3.  
DR PROSITE: PS50002; SH3; 1.  
KW SH3 domain; Repeat.  
FT DOMAIN 65 151 PDZ 1.  
FT DOMAIN 160 246 PDZ 2.  
FT DOMAIN 313 393 PDZ 3.  
FT DOMAIN 428 498 SH3.  
FT DOMAIN 534 724 GUANYLATE KINASE.  
SQ SEQUENCE 724 AA: 80472 MW: 78969.91 PF090BA CRC64:

Query Match 7.5%; Score 382; DB 1; Length 724;  
Best Local Similarity 26.2%; Pred. No. 1.5e-13;  
Matches 128; Conservative 74; Mismatches 143; Indels 144; Gaps 17;

OY 533 PESINTSSSTPAFAFASPARCP---IIPGCEITIE-----ISKRGIGLGLSTVGSD 583  
DB 26 PAHLNQA--NSPPVIVNTDTLEAPGELOVNGTEGEMEYEITLLENGSLGFSIAGTD 84  
OY 584 T-LG--AFIHEVEGAACKDGRMLAGDQILEVNGIDLRKATHEAIVNLRQTPQAV 639  
DB 85 NHIDDDPSIFITKIIIPGAAADGRRLRYNDISILFVNEVDVETHSNAVELKENGSLV 144  
OY 640 RLTYRDEAPYKEEYVCDTLTLELQKKPKGLGLSTVGK-----RNDGFEVSDIVKG 693  
DB 145 RLYVMRRKPAEK-----IIEIKLIKGP-KGLGFSIAGGVGNQHIIPGDSIYVTKIIEGG 198  
OY 694 IADPGRLQGOIILLNVEDVRNMQEVAALLKCSLGTVLEVCRIAGFHSRRS 753  
DB 199 AAHKGRLOIGKILAVNSVGLVDVHEDAVALKNTYDVYLKVA-----KPS 247  
OY 754 QTSQVSEGLSFTPLSGSTSESSSKKNALASEIOGLRTVEMKGPIDSLGISIA 813  
DB 248 -NAVLS--SYAPPDITTSYQHLDNEISHS----- 276  
OY 814 GGVSPGLGDVPFIAMHPTGYAAOTOKLRVDRIVTICGTSTEGMTHTQAVNLKNASG 873  
DB 277 -----YLGTDYPT-----AMTPT----- 289  
OY 874 SIEMQVAGGVSVYTGHHQPASSSLSTGLSTSIIPDDGLGPPCKITLERGPDGIG 933  
DB 290 -----SPRRSPYAKDLGLEDI-PREPRIVIHKGSTG 324  
OY 934 FSIVGGYSGPHGDLPIYKVFKAASDGRKRGDOIIAANGSLSEVTHEEVAAILK 993  
DB 325 FNIIVG---EDGE-GIFISFILAGPADLSCGLRKQDQILSYNGVDLRNASHQDAIALK 380  
OY 994 RTKGVTTLM 1002  
DB 381 NAGQVTTII 389

RESULT 10  
DLG1\_DROME STANDARD; PRT; 960 AA.  
ID DLG1\_DROME  
AC P31007;

DI 01-JUL-1993 (Rel. 26, Created)  
DI 01-JUL-1993 (Rel. 26, Last sequence update)  
DI 30-MAY-2000 (Rel. 39, Last annotation update)  
DE LETHAL(1)DISCS LARGE-1 TUMOR SUPPRESSOR PROTEIN.  
GN DLG1.

OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Eukaryota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
RX NCBI\_TaxID=7227;

RP SEQUENCE FROM N.A.  
RX MEDLINE=91330294; PubMed=1651169;  
RA Woods D.F., Bryant P.J.;  
RT "The discs-large tumor suppressor gene of Drosophila encodes a  
RT guanylate kinase homolog localized at septate junctions.";  
RL Cell 66:451-464(1991).

CC -1- FUNCTION: LOSS OF THIS PROTEIN CAUSES LARGE IMAGINAL DISKS BY  
CC ALLOWING NEOPLASTIC OVERGROWTH. THE PRESENCE OF A GUANYLATE KINASE  
CC DOMAIN SUGGESTS INVOLVEMENT IN ADEQUATE CELLULAR ADHESION AS WELL  
CC AS SIGNAL TRANSDUCTION TO CONTROL CELLULAR PROLIFERATION. THE DLG  
CC PROTEIN PLAYS A CRITICAL ROLE AT SEPTATE JUNCTIONS IN CELLULAR  
CC GROWTH CONTROL DURING LARVAL DEVELOPMENT.  
CC -1- SUBCELLULAR LOCATION: CYTOSKELETON-ASSOCIATED. LOCATED AT THE  
CC CYTOPLASMIC FACE OF THE MEMBRANE IN THE CELLULAR BLASTODERM AND  
CC BECOMES ASSOCIATED WITH SEPTATE JUNCTIONS WHICH BEGIN TO FORM  
CC BETWEEN EPITHELIAL CELLS AT THE TIME OF DORSAL CLOSURE. IN ADULT  
CC FLIES, LOCATED AT THE APICAL-LATERAL MEMBRANE BOUNDARY OF  
CC EPITHELIAL CELLS.

CC -1- ALTERNATIVE PRODUCTS: LEADS TO DIFFERENT FORMS OF THE PROTEIN;  
CC THE SEQUENCE SHOWN REFERS TO THE DLG-A PROTEIN, THAT IS MOST  
CC ABUNDANTLY EXPRESSED.

CC -1- TISSUE SPECIFICITY: BROAD SPECIFICITY. ABUNDANTLY EXPRESSED IN  
CC EPITHELIAL TISSUES. DEVELOPMENTAL STAGE DEPENDENT EXPRESSION.

CC -1- SIMILARITY: CONTAINS 3 PDZ/DRH DOMAINS.  
CC -1- SIMILARITY: CONTAINS A GUANYLATE KINASE-LIKE DOMAIN.  
CC -1- SIMILARITY: BELONGS TO THE MAGUK FAMILY OF CELL JUNCTION PROTEINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: M73529; AAA28468.1; -  
DR PIR: A39651; A39651.  
DR HSSP: Q12959; 1PDR.  
DR Flybase: FBgn0001624; dlgl.  
DR InterPro: IPR000619; -  
DR InterPro: IPR001452; -  
DR InterPro: IPR001478; -  
DR Pfam: PF00625; Guanylate\_kin; 1.  
DR Pfam: PF00595; PDZ; 3.  
DR Pfam: PF00018; SH3; 1.  
DR PROSITE: PS00856; GUANYLATE\_KINASE\_1; 1.  
DR PROSITE: PS50052; GUANYLATE\_KINASE\_2; 1.  
DR PROSITE: PS50106; PDZ; 3.  
DR PROSITE: PS50002; SH3; 1.  
KW Transducer; SH3 domain; Alternative splicing; Repeat.  
FT DOMAIN 40 126 PDZ 1.  
FT DOMAIN 154 244 PDZ 2.  
FT DOMAIN 486 566 PDZ 3.  
FT DOMAIN 600 670 SH3.  
FT DOMAIN 768 960 GUANYLATE KINASE.  
SQ SEQUENCE 960 AA: 102468 MW: 87874.4262 F1B6A5 CRC64:

Query Match 7.5%; Score 379; DB 1; Length 960;  
Best Local Similarity 24.8%; Pred. No. 3.2e-13;  
Matches 136; Conservative 85; Mismatches 190; Indels 138; Gaps 19;



[illegible]



DB 422 SSQSDNHVSPS-----SYLGQFPASPARSPIS 450  
 QY 835 VAAQFOKLRVDRIVITICSTEGMTHQAVNLKNASSIEMOVAVGVVTHHOE 894  
 DB 451 KAV-----LGDEEIT-----RE 462  
 QY 895 PASSSLFTGLNSTSIFODDLGPPCKSITLERGPDGLGFSIVGYGSPHGLPIYKTV 954  
 DB 463 P-----RKVYLHSGSTGLGPNIVG---EDGE-GIFISFI 493  
 QY 955 FAKGAASEDGRKRGDQIIAIVNGOSLEGVTHEEVAAILKRTKGTITLM 1002  
 DB 494 LAAGPADLSELGRKDRITISVNDLRAASHEDQAAALKNAGCAVITV 541  
 RESULT 12  
 DGL1\_HUMAN  
 ID DGL1\_HUMAN STANDARD; PRT; 904 AA.  
 AC Q12959; Q12958;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 01-OCT-2000 (Rel. 40, Last annotation update)  
 DE PRESYNAPTIC PROTEIN SAP97 (SYNAPSE-ASSOCIATED PROTEIN 97) (DISCS,  
 DE LARGE HOMOLOG 1).  
 GN DGL1.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 ON NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=95024052; PubMed=7937897;  
 RA Lue R.A., Marfatia S.M., Branton D., Chishti A.H.;  
 RT "Cloning and characterization of hdl9: the human homologue of the  
 RT Drosophila discs large tumor suppressor binds to protein 4.1.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 91:9818-9822(1994).  
 RN [2]  
 RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS) OF 460-555.  
 RX MEDLINE=96338231; PubMed=8757139;  
 RA Cabral J.H., Petosa C., Sutcliffe M.J., Raza S., Byron O., Poy F.,  
 RA Marfatia S.M., Chishti A.H., Liddington R.C.;  
 RT "Crystal structure of a PDZ domain.";  
 RL Nature 382:649-652(1996).  
 CC -1- FUNCTION: INTERACTS WITH THE CYTOPLASMIC TAIL OF NMDA RECEPTOR  
 CC SUBUNITS. ASSOCIATES WITH PROTEIN 4.1.  
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; 1 (SHOWN HERE) AND 2; ARE  
 CC PRODUCED BY ALTERNATIVE SPLICING.  
 CC -1- SIMILARITY: CONTAINS 3 PDZ/DRH DOMAINS.  
 CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.  
 CC -1- SIMILARITY: BELONGS TO THE MAGUK FAMILY OF CELL JUNCTION PROTEINS.  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC  
 CC EMBL: U13897; AAA50599.1; -  
 CC EMBL: U13896; AAA50598.1; -  
 CC PDB: 1PDR; 23-JUL-97.  
 CC MIM: 601014; -  
 CC InterPro: IPR000619; -  
 CC InterPro: IPR001452; -  
 CC InterPro: IPR001478; -  
 CC Pfam: PF00625; Guanylate\_kin; 1.  
 CC Pfam: PF00595; PDZ; 3.  
 CC Pfam: PF00018; SH3; 1.  
 CC PROSITE: PS00856; GUANYLATE\_KINASE\_1; 1.  
 CC PROSITE: PS0052; GUANYLATE\_KINASE\_2; 1.  
 CC PROSITE: PS0106; PDZ; 3.

DR PROSITE; PS00002; SH3; 1.  
 KW SH3 domain; Repeat; Alternative splicing; 3D-structure.  
 FT DOMAIN 224 310 PDZ 1.  
 FT DOMAIN 319 405 PDZ 2.  
 FT DOMAIN 466 546 PDZ 3.  
 FT DOMAIN 581 651 SH3.  
 FT DOMAIN 714 904 GUANYLATE KINASE.  
 FT VARSPIC 669 680 EIPDDPSKGLK -> OSFNDKRNKRLFRKPPFYKNDQS  
 FT SEQUENCE 904 AA; 100354 MM; B7879D6B0920D4 CRC64;  
 SQ  
 Query Match 7.4%; Score 377; DB 1; Length 904;  
 Best local similarity 29.9%; Pred. No. 3.8e-13;  
 Matches 119; Conservatve 58; Mismatches 129; Indels 92; Gaps 14;  
 QY 634 QTPQVRRLTYRDEAPYKEEVCDTLTIELKKPKGLSIVGR-----NDTGVS 687  
 DB 208 ETPTYVNGT---DAVEYEI---TLE---RNGSLGFSINGTDNPHIGDSSIFIT 286  
 QY 688 DIYKGGIADPDGRLIGDDQILLVNGEDVYRNASQEAVALKCSLGYTTEVGRKAGPFH 747  
 DB 257 KIITGGAADGRLRYNDCLIQVNEVDYRDVTSKAVEALKEGSIYRLVYKR----- 309  
 QY 748 SEHRPSQTSQVSEGLSSEFFPLSGSTSESSSKNALASEIOLGRTVEKKGPTDS 807  
 DB 310 --RKP-----VSE-----KIMEIKLIKGP-KG 328  
 QY 808 LGISIAGVSP--LGDPVPIFAMHPYGAOTOKLRVDRIVITICSTEGMTHQAV 865  
 DB 329 LGFSIAGVGNQIHIPGNSIYVKRIIEGAAHDKLQIGDKLLAVNNVLEEVTHEAV 388  
 QY 866 NLKNASGSIEMOVAVAGDVSV-----VTGHQEPAS---SLSFTGLT----- 906  
 DB 389 TALKNTSDFYLLKVAAPTSYKMDGYAPDITNNSSQPDNHVSPSFLQTPASPARYS 448  
 QY 907 --STSIPODLGPPCKSITLERGPDGLGFSIVGYGSPHGLPIYKTVFAKGAASEDG 964  
 DB 449 PVKAVAYGDEETREPRKVVYLHSGSTGLGPNIVG---EDGE-GIFISFILAGPADLSG 504  
 QY 965 RLKRGDQIIAIVNGOSLEGVTHEEVAAILKRTKGTITLM 1002  
 DB 505 ELKRGDRITISVNDLRAASHEDQAAALKNAGCAVITV 542  
 RESULT 13  
 IL16\_HUMAN  
 ID IL16\_HUMAN STANDARD; PRT; 631 AA.  
 AC Q14005; Q16435; Q9UP18;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE INTERLEUKIN-16 PRECURSOR (IL-16) (LYMPHOCYTE CHEMOTACTRANT FACTOR)  
 DE (LCEP).  
 GN IL16.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 ON NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=peripheral blood;  
 RA Kornfeld H.;  
 RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99145586; PubMed=9990060;  
 RA Bannert N., Avots A., Baler M., Serfling E., Kurth R.;  
 RT "GA-binding protein factors, in concert with the coactivator CREB  
 RT binding protein/p300, control the induction of the Interleukin 16  
 RT promoter in T lymphocytes.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 96:1541-1546(1999).  
 RN [3]

RP SEQUENCE OF 241-631 FROM N.A.  
 RA MEDLINE-96211141; PubMed-8609984;  
 RA Bannert N., Baier M., Werner A., Kurth R.;  
 RT "Interleukin-16 or not?";  
 RL Nature 381:30-30(1996).  
 RN [4]  
 RP SEQUENCE OF 502-631 FROM N.A.  
 RC Tissue-Peripheral blood;  
 RX MEDLINE-94255480; PubMed-7910967;  
 RA Crikshank W.W., Center D.M., Nisar N., Wu M., Natke B.C.,  
 RA Theodore A.C., Kornfeld H.;  
 RT "Molecular and functional analysis of a lymphocyte chemoattractant  
 factor: association of biologic function with CD4 expression.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 91:5109-5113(1994).  
 RN [5]  
 RP SEQUENCE OF 502-631 FROM N.A.  
 RA Du Y., Du G.X., Hou L.H., Wang H.T.;  
 RT "CDNA sequence of Interleukin-16 cloned from human PBLS.";  
 RL Submitted (MAR-1998) to the EMBL/Genbank/DBJ databases.  
 CC -1- FUNCTION: STIMULATES A MIGRATORY RESPONSE IN CD4+ LYMPHOCYTES,  
 CC MONOCYTES, AND EOSINOPHILS. ALSO INDUCES T-LYMPHOCYTE EXPRESSION  
 CC OF INTERLEUKIN 2 RECEPTOR. LIGAND FOR CD4.  
 CC -1- SUBUNIT: HOMOTETRAMER.  
 CC -1- SUBCELLULAR LOCATION: SECRETED.  
 CC -1- SIMILARITY: CONTAINS 2 PDZ/DHR DOMAINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial/  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: M90391; AAD04636.1; -  
 DR EMBL: M9077011; AAD15990.1; -  
 DR EMBL: S81601; AAB36371.1; -  
 DR EMBL: AF053412; AAC12732.1; -  
 DR MIM: 603035; -  
 DR HSSP: Q12959; 1PDR.  
 DR InterPro: IPR001478; -  
 DR Pfam: PF00595; PDZ 1.  
 DR PROSITE: PS0106; PDZ 2.  
 KW Cytokine; Chemotaxis.  
 FT PROPEP 1 631  
 FT CHAIN 1  
 FT DOMAIN 411 496 PDZ 1.  
 FT DOMAIN 533 618 PDZ 2.  
 FT CONFLICT 104 104 E -> D (IN REF. 2).  
 FT CONFLICT 233 233 F -> L (IN REF. 2).  
 FT CONFLICT 241 243 LRL -> PRE (IN REF. 3).  
 FT CONFLICT 319 319 E -> A (IN REF. 1).  
 SO SEQUENCE 631 AA; 66694 MW; 2A9CED70CC8D4AC CRC64;

Query Match 5.68; Score 283; DB 1; Length 631;  
 Best Local Similarity 22.0%; Pred. NO. 2.8e-08;  
 Matches 171; Conservative 106; Mismatches 306; Indels 194; Gaps 31;

DB 4 SFDTPAEPWVRISPCINLNLSPISEB-----GHPLOPNASLNEEGTQG----- 51  
 OY 316 DPNGAAGRGRLQADELLEINGQLYGRSHONA-SSTIKCAPS-----KVIIFIR 366  
 DB 52 HPDGRPR-----LDTANGT---PKYKASDSSIVKGPYPAPKAPRROSLKGLR 99  
 OY 367 NKDAVNAQVCPGNAVEPLPSNSENLQKETEPTVTSDAVAVDSSFKNVQHLLEPKDQG 426  
 DB 100 NRASPRGLPPLALSTGCPAPASREHL-GSHIRASSSSSSIRORISSFETFGSSQLP-DKG 157  
 OY 427 GLGTAISEDLSTGYITISLTHGVAATDGRKAGDQILAVDELIVGPIEKFISLTKT 486  
 DB 158 AQRSLQPS---SGEAPKPLGNH-----EGR-----FSGLT-- 186  
 OY 487 AKMTKLTIHENPQSAVPAAGAAAGEKKNSSQSLWVQSGSPESISINTSSSTPA 546  
 DB 187 -----GRGNA-----PTLVPO---QPRQV---LSSGSPA 210  
 OY 547 IFASDPATCPIPGCETTYEISKGRGGLSLVSGSDPLTAFIHEVEGACAKDQRL 606  
 DB 211 SEARPGVSESP-----PGRQPNQKTFPPPPDPLRLSLQARESGGP----- 254  
 OY 607 WAGDQILEVNGIDLRKATHDEAINVLRQTPQVRVLTLYRDEAPYKKEEVCOTLTLEOK 666  
 DB 255 -----VLKMPQORAR-----SPPLTRSSCETKL--LDEK 282  
 OY 667 PGKGLSLVGRKNDTGVFSDIVYKGIADPDGRILQGDQILVLVNGEDVRNASQPAVAL 726  
 DB 283 TSKLYSIS-----SOVSSAVKMSLCLDPSSISCAOTPCIPKGAAPTSSNEDSA- 333  
 OY 727 LKCSLGTVTLVYGRKAGPFH---SERP-----SQTQVSEGLSLSPFPLSGSTSES 778  
 DB 334 -NGSAETSAIDTG-----FSLNISELREYEGLEAKEDDGDHSSIQ---SGQSVISL 383  
 OY 779 LESSSK-----NALSEIOGLTVEKKGPTSLGSLINGVSGSLGDPVPIIA 828  
 DB 384 LSSEELKRLIEVKYLDATLQDLGIVHTLILKEGAGLSESLAG--GADLENKVTIVH 441  
 OY 829 MNHPTGVAQOTOKLRVGRIVTICSTEGMTHQAVMLKNASGSIEMQVYAGDSVY 888  
 DB 442 RYFPRGLASQSGTQKNGEVLSTNGSKSLKGTTHHALILNQAARPRQAVYTRLRPEA 501  
 OY 889 TGHQEPASSLSIFGLTSTSPDQDLPPPOCKSTTLERGDGLGFSIVGYGSHGDL 948  
 DB 502 MPDLNSTSDSAASASMAADSV--ESTAEAVYCVTLTKMSAGLQFSLGKGSJHGDKP 559  
 OY 949 IYKTVFAKGAASEDGR-LKRGDQIIAVNGSLSEVTHEEVAIKR-TKGVTLMV 1003  
 DB 560 LTINRIF-KGAASEQSEVQPODEILQGLTRAGSLTFEAMNITKALPDGPVITI 615

RESULT 14  
 ZOI\_HUMAN  
 ID ZOI\_HUMAN STANDARD; PRT: 1736 AA.  
 AC 007157;  
 DT 01-OCT-1994 (Rel. 30, Created)  
 DT 01-OCT-1994 (Rel. 30, Last sequence update)  
 DT 01-OCT-2000 (Rel. 40, Last annotation update)  
 DE TIGHT JUNCTION PROTEIN ZOI-1 (ZONA OCCUDENS 1 PROTEIN) (TIGHT JUNCTION  
 DE PROTEIN 1).  
 GN TUP1 OR ZOI.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Euthera; Primates; Catarrhini; Hominoidea; Homo.  
 OX NCBI\_Taxid=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC Tissue-Liver;  
 RX MEDLINE-93361541; PubMed-8395056;  
 RA Willott E., Balda M.S., Fanning A.S., Jameson B., van Itallie C.,  
 RA Anderson J.M.;  
 RT "The tight junction protein ZO-1 is homologous to the Drosophila  
 RT discs-large tumor suppressor protein of septate junctions.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 90:7834-7838(1993).  
 CC -1- FUNCTION: THE N-TERMINAL MAY BE INVOLVED IN TRANSDUCING A SIGNAL  
 CC REQUIRED FOR TIGHT JUNCTION ASSEMBLY, WHILE THE C-TERMINAL MAY  
 CC HAVE SPECIFIC PROPERTIES OF TIGHT JUNCTIONS. THE ALPHA DOMAIN  
 CC MIGHT BE INVOLVED IN STABILIZING JUNCTIONS.  
 CC -1- SUBUNIT: INTERACTS WITH OCCUDIN, CLAUDINS AND ZO-3.  
 CC -1- SUBCELLULAR LOCATION: PERIPHERAL MEMBRANE, CYTOPLASMIC SIDE.  
 CC MOVEMENT OF ZO-1 FROM THE CYTOPLASM TO MEMBRANE IS AN EARLY EVENT  
 CC OCCURRING CONCURRENTLY WITH CELL-CELL CONTACT.  
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS: A LONG FORM (SHOWN HERE) AND A  
 CC SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.  
 CC -1- TISSUE SPECIFICITY: THE ALPHA-CONTAINING ISOFORM IS FOUND IN MOST

```

Epithelial Cell Junctions, The Short Isoform Is Found Both In
Endothelial Cells And The Highly Specialized Epithelial Junctions
Of Renal Glomeruli And Sertoli Cells Of The Seminiferous Tubules.
-1- PPM: PHOSPHORYLATED.
-1- SIMILARITY: CONTAINS 3 PDZ/DB DOMAINS.
-1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
-1- SIMILARITY: BELONGS TO THE MAGUK FAMILY OF CELL JUNCTION PROTEINS.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
DR EMBL: L14837; AAA02891.1; -.
DR PIR: A47747; A47747.
DR HSP: Q12959; IPDR.
DR MIM: 601009; -.
DR InterPro: IPR000619; -.
DR InterPro: IPR000906; -.
DR InterPro: IPR001452; -.
DR InterPro: IPR001478; -.
DR Pfam: PF00623; Guanylate_kin; 1.
DR Pfam: PF00595; PDZ; 3.
DR Pfam: PF00018; SH3; 1.
DR Pfam: PF00791; SH3; 1.
DR PROSITE: PS00856; GUANYLATE_KINASE_1; FALSE_NEG.
DR PROSITE: PS50052; GUANYLATE_KINASE_2; 1.
DR PROSITE: PS50106; PDZ; 3.
DR PROSITE: PS50002; SH3; 1.
KW Tight junction; SH3 domain; Alternative splicing; Repeat; Membrane;
KW Phosphorylation.
FT DOMAIN 11 98 PDZ 1.
FT DOMAIN 174 252 PDZ 2.
FT DOMAIN 409 490 PDZ 3.
FT DOMAIN 504 572 SH3.
FT DOMAIN 632 782 GUANYLATE_KINASE.
FT DOMAIN 1231 1236 POLY-PRO.
FT DOMAIN 1414 1420 POLY-PRO.
FT VASPLIC 910 989 MISSING (IN SHORT ISOFORM).
SQ SEQUENCE 1736 AA; 194721 MW; 508D01B7A0814FFE CRC64;

Query Match 4.8%; Score 243; DB 1: Length 1736;
Best Local Similarity 20.0%; Pred. No. 1.8e-05;
Matches 100; Conservative 68; Mismatches 149; Indels 182; Gaps 15;

QY 651 KEEVCDLTLTTELKPKPKGLSLIVGKRN-----DTGVSDIVKGIADPGRLIQ 703
: : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2 EETAWMEQHTYTLHNAPEFGFICIALSGGRDNHPSGELSIYISDVLYKGPA--EGQLDE 59.
QY 704 GDQIILVNGEDVRNASQE-AVAALLKCSLGT-----VTELVGRIKAGFHSERRPS 753
: : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 60 NDRVAMVNGVSDMNEHAFAVOQLKRSKNAKITIRKKKQVQIPVSRPDEPVSDNEEDS 119
QY 754 QTSOVSESSLSFTPLPG-----STSSLESSSKKNLAQE--IQ 793
: : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 120 YDEELHD-----PRSGRGVGNVRSEKITWPRDRSASRERSLSPRSRRRVASSOPAK 171
QY 794 GLRVEKKGPDLISGLISAGVGSPLDVPIFIAMHPTGVAAQTKLRVGDRLVITCG 853
: : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 172 PTKYTLVSKRNEEGLRLAS-----HIFVEISODSLAANDGNIQEGDVYLYKNG 222
QY 854 TSTEGMTHTQAVNLKNAAGSIEMQVAG-----GDVSV 887
: : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 223 TVTEWMSLTDAKTLIERSKGLKMYVQDERATILNVPLDLSHSANASERDDIETQS 282
QY 888 VTGHH-----QEPASSL-----SFGTLNST 908
: : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 283 LASDHSGRSHDRPPRRSRSPDQRSEPSDHSRHSPPQPSNGLSRDEERISKGAVST 342

```

```

OY 909 SI+PD0DLGPPOCKSITLERGP----- 929
Db 343 PVKADHDPTVEVEVYERNEKOTPSLPDEKPPVYAQVGNOMIYLSVHLNVSTLIQMLK 402
QY 930 -----DELGFSIVGYGSPGHDLPITYKTVFPANGASEDRLRKQDII 973
Db 403 MGLFRPSMKLVKFRKGSVGLRLAG-----NDVGIFAVGLDESPAKEG-LREGDPII 456
QY 974 AVNGQSEGVTHEEAVIL 992
Db 457 RVNNVDFNIREEAVLFL 475

RESULT 15
ID_ZO3_HUMAN STANDARD: PRT; 933 AA.
AC 095049;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE TIGHT JUNCTION PROTEIN ZO-3 (ZONA OCCUDENS 3 PROTEIN) (TIGHT JUNCTION
DE PROTEIN 3).
GN TJP3 OR ZO3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN (1)
RA Lamerdin J.E., McCreedy P.M., Skowronski E., Viswanathan V.,
RA Burkhardt-Schultz K., Brower A., Gordon L., Dias J., Ramirez M.,
RA Stillegen S., Phan H., Velasco N., Do L., Regala W., Terry A.,
RA Ganes J., Dangnan L., Erler A., Christensen M., Georgescu A.,
RA Avila J., Liu S., Attix C., Andreise T., Frankheim M.,
RA Amico-Keller G., Coefield J., Duarte S., Lucas S., Bruce R.,
RA Thomas P., Quan G., Kronmiller B., Arellano A., Sanders C., Ow D.,
RA Nolan M., Trong S., Kobayashi A., Olsen A.S., Carraro A.V.;
RT "Sequence analysis of a 3.5 kb contig in human 19p13.3 containing a
RT serine protease gene cluster.";
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
CC -1 SUBUNIT: INTERACTS WITH OCCUDIN, CLAUDIN AND ZO-1 (BY
CC SIMILARITY).
CC -1 SIMILARITY: CONTAINS 3 PDZ/DHR DOMAINS.
CC -1 SIMILARITY: CONTAINS 1 SH3 DOMAIN.
CC -1 SIMILARITY: CONTAINS A GUANYLATE KINASE-LIKE DOMAIN.
CC -1 SIMILARITY: BELONGS TO THE MAGUK FAMILY OF CELL JUNCTION PROTEINS.
CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR DR EMBL: AC005954; AAC72274.1; ALT_INIT.
DR HSP: P31016; 18PF.
DR InterPro: IPR000619; -
DR InterPro: IPR001478; -
DR InterPro: IPR001478; -
DR Pfam: PF00595; PDZ; 3.
DR Pfam: PF00625; Guanylate_kin; 1.
DR PROSITE: PS00856; GUANYLATE_KINASE_1; FALSE_NEG.
DR PROSITE: PS50052; GUANYLATE_KINASE_2; 1.
DR PROSITE: PS50002; SH3; 1.
DR PROSITE: PS50106; PDZ; 3.
DR Tight junction; SH3 domain; Repeat; Membrane.
FT DOMAIN 11 93 PDZ 1.
FT DOMAIN 195 272 PDZ 2.
FT DOMAIN 394 460 PDZ 3.
FT DOMAIN 489 563 SH3.
FT DOMAIN 675 775 GUANYLATE KINASE.

```

SEQUENCE 933 AA: 102800 MW: 680298CFD0615B47 CRC64:

Query Match 4.8%: Score 242; DB 1: Length 933;

Best Local Similarity 22.8%: Pred. No. 8.2e-06; Matches 121; Conservative 62; Mismatches 155; Indels 192; Gaps 20;

```
QY 564 TIEISKG-RTGLGSIYVGSPTLGAFLIHEVEGAACKDGLMAGDQILEVNGIDLK 622
Db 11 TATLSKDPDRRGFGIAISGGRDPGSMVSDVPGGPA-EGRLQTGDHIYMWNGVSMEN 68
QY 623 ATHDEATVLR-----QTPOVRLTLXR-----DEAPYKEEYV----- 655
Db 69 ATSAFAIQILKTCMANITVRRRIHLPATKASPSSPGRQSDSDDDGPQVEEVDQGR 128
QY 656 -----CDTLTI-----EL 663
Db 129 GYDDSSSGSGSRNDESRRRPRGRAGSHGRSPGGSGSEANGALALVSGFKRLPRQDV 188
QY 664 QKKPGKGLGSIYVKRND-----TGVEVSDIVKGIADPDGRLIQDQIILLVNGED 714
Db 189 QMKPVK---SVLVKRRDSEFEGVKGSIQIFIKHITDGLAARHRLQEGDLILQINGVS 244
QY 715 VRNASQEAVALLKCSLCTVTLVGR-----740
Db 245 SONSLNDTRRLIEKSEGLSLVLVDRGQFLVNIIPPAVSDSDSSPLEEGVTMADEMSSP 304
QY 741 -----IKAGPFHSERRP--SOTSQY-----SEGLSSTFEPPLSGS 773
Db 305 PADISDLASELSQAPPSHIIPPRHAKRSPASQTDSPVESPRLRRESSVDSRTISEPDE 364
QY 774 STSESLESSS-----KKNALASEIQG---LRTYEMKKGPTDSLGIISTAGVGSPPLGDPV 824
Db 365 QRSELPRESSYDIYRVPSQSMEDRGYSPDTRVYVFLKG--KSIGLRLAGG-----NDVG 417
QY 825 IEIIMMHTGYAAQTKRVGDRIVTICPTSEGTHTQAVNLKKNAGSIEMQVYAG-- 882
Db 418 IFVSGVQ-AGSPADGGQIQEGDQILQVNDVPFQNLTRREAVQFLGLGPPGGEEMELVYORK 476
QY 883 -----GDVSVVTGHHQ-EPASSLSFTGLTSTSIQO--DDLGP 917
Db 477 QDIFMKVQSRVGDSTFYIRTHFELEP--SPPSGLGFTRGDVFHVDITLHP 524
```

Search completed: July 12, 2001, 14:43:25  
Job time: 218 sec



```

QY 181 LSNGEVNRGIFIKHVLKEDSPAGKNGTLKPGDRIVE-----215
DB 1176 LSNGEVNRGIFIKHVLKEDSPAGKNGTLKPGDRIVEVGDMDLRDASHQAVAIKRNKPNV 1235
QY 216 -----ASQSESEPEKAPLC 230
DB 1236 VFMVQSIINRRKSPPLSLHLNLYPKYNSSTNPFADSLQIMDKAPSSQSESEPEKAPLC 1295
QY 231 SVPPPPSARAEMGSDHTOSSASKISODYDKEDEFGYSWMKINRERYCTLGELMIELEK 290
DB 1296 SVPPPPSARAEMGSDHTOSSASKISODVDKEDDFGYSWMKINRERYCTLGELMIELEK 1335
QY 291 GHSGLSGLSLAGNDRSMYSVFIYVIGIDPNGAAGKGRLOIADELLEINGOILLYGRSHONAS 350
DB 1356 GHSGLSGLSLAGNDRSMYSVFIYVIGIDPNGAAGKGRLOIADELLEINGOILLYGRSHONAS 1415
QY 351 SIICAPSKVKIIFIRNKDAVNQMAVCPGNAVEPLPSNSENLOKKEPEPTVTSDAAYDL 410
DB 1416 SIICAPSKVKIIFIRNKDAVNQMAVCPGNAVEPLPSNSENLOKKEPEPTVTSDAAYDL 1475
QY 411 SSFKNVQHLKPKDGGGLGIAISEEDTLGVIITKSLTEHGVAATDGRLLKYGDOILAVDDE 470
DB 1476 SSFKNVQHLKPKDGGGLGIAISEEDTLGVIITKSLTEHGVAATDGRLLKYGDOILAVDDE 1535
QY 471 IVVGYPIEKFIISLKTAKMTVTKLTHAENPDQAVPSAAGASGEKKNSOSLWVPOSGS 530
DB 1536 IVVGYPIEKFIISLKTAKMTVTKLTHAENPDQAVPSAAGASGEKKNSOSLWVPOSGS 1595
QY 531 PEPBSINTSRSSSTPAIFASDPATCPIIPGCEETIETISKRTGLGSLIVGSDTLGCAFI 590
DB 1596 PEPBSINTSRSSSTPAIFASDPATCPIIPGCEETIETISKRTGLGSLIVGSDTLGCAFI 1655
QY 591 IHEVYEEGAAKDKGRRLMAGQOILEVNGIDLRKATHDEAINVLRQTPQVRLLTLRDEAPY 650
DB 1656 IHEVYEEGAAKDKGRRLMAGQOILEVNGIDLRKATHDEAINVLRQTPQVRLLTLRDEAPY 1715
QY 651 KEEBVCDTLTLELQKPKGKGLGSLIVGKRNDYGVFVSDIYKGIADPDGRLLQGDQILLY 710
DB 1716 KEEBVCDTLTLELQKPKGKGLGSLIVGKRNDYGVFVSDIYKGIADPDGRLLQGDQILLY 1775
QY 711 NGEVNRASQENAVALLKCSLGYTTELEVGRITKACPFHSERRPSQTSQVSEGLSSFFPPL 770
DB 1776 NGEVNRASQENAVALLKCSLGYTTELEVGRITKACPFHSERRPSQTSQVSEGLSSFFPPL 1808
QY 771 SCSSTSELSSSSKKNALASEIOGLRTVEKKKGPDTSLGISIAGVGSPLGDPVFIETAM 830
DB 1809 -GSSSTSELSSSSKKNALASEIOGLRTVEKKKGPDTSLGISIAGVGSPLGDPVFIETAM 1867
QY 831 HPFGVAAQOTKLRVGRIVITICGTSTEGMTHTQAVNLLKNAAGSIEQVYVAGDVSVAVTG 890
DB 1868 HPFGVAAQOTKLRVGRIVITICGTSTEGMTHTQAVNLLKNAAGSIEQVYVAGDVSVAVTG 1927
QY 891 HHOEPASSISLFTGLTSTSTFODDLGPPROCKSTLLEGGPGLGSLIVGSGSPHGDLPY 950
DB 1928 HHOEPASSISLFTGLTSTSTFODDLGPPROCKSTLLEGGPGLGSLIVGSGSPHGDLPY 1987
QY 951 VKTVFAGAAASEDGRRLKRGDQIIAVNGQSLEGYTHEEVAVALIKRTKGTVTLMYLS 1005
DB 1988 VKTVFAGAAASEDGRRLKRGDQIIAVNGQSLEGYTHEEVAVALIKRTKGTVTLMYLS 2042

```

```

RESULT 2
OS5164 ID 055164 PRELIMINARY: PRT: 2054 AA.
AC 055164:
DT 01-JUN-1998 (Tremblrel. 06, Created)
DT 01-JUN-1998 (Tremblrel. 06, Last sequence update)
DE 01-MAR-2001 (Tremblrel. 16, last annotation update)
DE MULTIPDZ DOMAIN PROTEIN 1.
GN MUPPL.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

```

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
RX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RX MEDLINE=98196865; PubMed=9537516;
RA Ulmer C., Schuck K., Flage A., Lubbert H.,
RT "Cloning and characterization of MUPPL, a novel PDZ domain protein."
RL FEBS Lett. 424:63-68(1998).
DR EMBL: AJ001320; CA004681.1;
DR HSSP: Q12959; 1PDR
DR InterPro: IPR001478;
DR Pfam: PF00595; PDZ; 13.
DR SMART: SM00228; PDZ; 1.
SO SEQUENCE 2054 AA; 218590 MW; 44BD3F42B801F78F CRC64;

```

```

Query Match 86.8%; Score 4413.5; DB 11; Length 2054;
Best local similarity 81.5%; Pred. No. 5.1e-258;
Matches 875; Conservative 64; Mismatches 62; Indels 73; Gaps 2;

```

```

QY 2 LQVSKSEPERTINIAKNSLIGMTVSAKNDGLGMYRSTIHGSAISRDRVAGDCLIS 61
DB 984 LQVSKSEPERTINIAKNSLIGMTVSAKNDGLGMYRSTIHGSAISRDRVAGDCLIS 1043
QY 62 INESTISVTNAQARALRRHSLIGPDIKITVYPAEHLKEFKISLGOOSGRVALDIFSS 121
DB 1044 INESTISVTNAQARALRRHSLIGPDIKITVYPAEHLKEFRVSVFGQAGIAMDIFSS 1103
QY 122 YTGADIDELPEBEEGESEELONTAVSNNOPRVELMRBPSKGLSIVGKMGSRLL 181
DB 1104 YTGADIDELPEBEEGESEELONTAVSNNOPRVELMRBPSKGLSIVGKMGSRLL 1163
QY 182 SNGEVMRGIFIKHVLKEDSPAGKNGTLKPGDRIVE-----215
DB 1164 SNGEVMRGIFIKHVLKEDSPAGKNGTLKPGDRIVEVGDMDLRDASHQAVAIKRNKPNV 1223
QY 216 -----APQSESEPEKAPLC 231
DB 1224 FVMQSIINRRKSPPLSLHLNLYPKYNSSTNPFADSLQIMDKAPSSQSESEPEKAPLC 1283
QY 232 VPPPPSARAEMGSDHTOSSASKISODVDKEDDFGYSWMKINRERYCTLGELMIELEK 291
DB 1284 VPPPPSARAEMGSDHTOSSASKISODVDKEDDFGYSWMKINRERYCTLGELMIELEK 1343
QY 292 HSGLSGLSLAGNDRSMYSVFIYVIGIDPNGAAGKGRLOIADELLEINGOILLYGRSHONAS 351
DB 1344 HSGLSGLSLAGNDRSMYSVFIYVIGIDPNGAAGKGRLOIADELLEINGOILLYGRSHONAS 1403
QY 352 SIICAPSKVKIIFIRNKDAVNQMAVCPGNAVEPLPSNSENLOKKEPEPTVTSDAAYDL 411
DB 1404 SIICAPSKVKIIFIRNKDAVNQMAVCPGNAVEPLPSNSENLOKKEPEPTVTSDAAYDL 1463
QY 412 SSFKNVQHLKPKDGGGLGIAISEEDTLGVIITKSLTEHGVAATDGRLLKYGDOILAVDDE 471
DB 1464 SSFKNVQHLKPKDGGGLGIAISEEDTLGVIITKSLTEHGVAATDGRLLKYGDOILAVDDE 1523
QY 472 IVVGYPIEKFIISLKTAKMTVTKLTHAENPDQAVPSAAGASGEKKNSOSLWVPOSGS 531
DB 1524 IVVGYPIEKFIISLKTAKMTVTKLTHAENPDQAVPSAAGASGEKKNSOSLWVPOSGS 1580
QY 531 EPBSINTSRSSSTPAIFASDPATCPIIPGCEETIETISKRTGLGSLIVGSDTLGCAFI 591
DB 1581 EPBSINTSRSSSTPAIFASDPATCPIIPGCEETIETISKRTGLGSLIVGSDTLGCAFI 1640
QY 592 HEVYEEGAAKDKGRRLMAGQOILEVNGIDLRKATHDEAINVLRQTPQVRLLTLRDEAPY 651
DB 1641 HEVYEEGAAKDKGRRLMAGQOILEVNGIDLRKATHDEAINVLRQTPQVRLLTLRDEAPY 1700
QY 652 EEBVCDTLTLELQKPKGKGLGSLIVGKRNDYGVFVSDIYKGIADPDGRLLQGDQILLY 711
DB 1701 EEBVCDTLTLELQKPKGKGLGSLIVGKRNDYGVFVSDIYKGIADPDGRLLQGDQILLY 1760

```

|    |      |  |      |
|----|------|--|------|
| Qy | 712  | GEDVRNNAQOEVAALLKCSIGTWTLEVGRKAKPFISSERPSSOTSOVSQSGISLSTFPLS | 771  |
| Db | 1761 | GEDVRNNAQOEVAALLKCSIGTWTLEVGRKAKPFISSERPSSOSOVSSLSLSTFPLS    | 1822 |
| Qy | 772  | GSSTSESSLESSKKNALASEIOGLRTVEEMKGPETDISIGISIAQVSPGLDVPITFAMH  | 831  |
| Db | 1831 | GIHTSESSSESSAKKNALASEIOGLRTVEIKGSPADALGISIAQVSPGLDVPITFAMH   | 1880 |
| Qy | 832  | PTGVAAQOKRKYVDRLVYTTCSTSEBMTQAVNLKNASSIEQVYVAGGDVSYVGH       | 891  |
| Db | 1881 | PNGVAAQOKRKYVDRLVYTTCSTSEBMTQAVNLKNASSIEQVYVAGGDVSYVGH       | 1940 |
| Qy | 892  | HOEPASSLSITFGLTSTISIFODDLGPPQCKSITLERGPOSLGRSIVYGSPHGLPIYV   | 951  |
| Db | 1941 | QOELANPCLAFETGLTSTSIFFDDLGPPQKSTITLDRGPOSLGRSIVYGSPHGLPIYV   | 2000 |
| Qy | 952  | KTVYPAKGAASDGLKRGDDOIAIANGSLGCVYTHEEVAVALKTKCTVTLAMLS        | 1005 |
| Db | 2001 | KTVYPAKGAALDGLKRGDDOIAIANGSLGCVYTHEEVAVALKTKCTVTLAMLS        | 2054 |

```

RESULT      3
0921K3
ID      0921K3      PRELIMINARY;      PRT: 2055 AA.
AC      0921K3;
DT      01-MAY-1999 (TREMBLrel. 10, Created)
DT      01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT      01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE      MULTIPLE PDZ DOMAIN PROTEIN.
        MPDZ.
GN      Mus musculus (Mouse).
OS      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Euteria; Rodentia; Sciurognathi; Muridae; Murinae; Mus
OX      NCBI_TaxID=10090;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=C57/BL6 X CBA F1; TISSUE=WHOLE BRAIN;
RA      Simpson E.H., Suffolk R., Jackson I.J.;
RT      "Identification and mapping of mouse Multiple PDZ domain protein,
RT      Mpdz."
RL      Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
DR      EMBL; AJ131869; CAA10523.1; -.
DR      HSSP; Q12959; 1PDR.
DR      MGD; MGI:1343489; Mpdz.
DR      InterPro; IPR001478; -.
DR      Pfam; PF00595; PDZ; 13.
DR      SMART; SM00228; PDZ; 1.
SQ      SEQUENCE 2055 AA; 218968 MW; E1A38EE6E34CC20EA CRC64;

```

|                       |       |              |        |            |    |        |      |
|-----------------------|-------|--------------|--------|------------|----|--------|------|
| Query Match           | 84.8% | Score        | 4314.5 | DB         | 11 | Length | 2055 |
| Best Local Similarity | 79.7% | Pred. No.    | 5e-252 |            |    |        |      |
| Matches               | 856   | Conservative | 71     | Mismatches | 74 | Indels | 73   |
|                       |       |              |        |            |    | Gaps   | 2    |

|    |      |   |      |
|----|------|---|------|
| OY | 2    | LONWKEEFERTINIAKNSNLSIGTVAANDGJMYRSTIHOGALSTRGRJAIGCIIS                   | 61   |
| Dd | 985  | LQSMQEAFERPRVITANGSSLSGMTTVA <del>SAL</del> KDGLGVYRSTIHOGALSTRGRJAVGCILS | 1044 |
| OY | 62   | INEESTISVTNAOARAMLRRHSLIGDPICKITYYPAAHELEBKTSLGQSGRYMALDIFFS              | 121  |
| Dd | 1045 | INEESTISLTNAQAARAMLRHSLSLGPDKITYYPAAHELEEFFRVFSFOQAAGIMADIFRS             | 1104 |
| OY | 122  | YTGRDIPPLPREAGEEGEESLONTAVSNMNOPRVEMPEPSKSIGISTVGGRMGSL                   | 181  |
| Dd | 1105 | YTGRDIPPLPREAGEEGEESLQNAAYSSMSOPRRVELMRPEKSLSIGISTVGGRMGSKRL              | 1166 |
| OY | 182  | SNGEVMRGCIFIKHVLEDSPACKNGTLKPGDRIVE-----                                  | 215  |
| Dd | 1165 | SNGEVMRGCIFIKHVLEDSPACKNGTLKPGDRITIEVDOMDLRDASHBQAVEAIRKANPYY             | 1224 |
| OY | 216  | -----APSGESEPERAPLCS  | 231  |
| Dd | 1225 | FMYOSIINRPKKSPLPHSLPYPKYSFSSSTNPADSLQTTDOAQSDBSESTEKFPALCN                | 1284 |

|    |      |   |       |
|----|------|---|-------|
| Qy | 232  | VPPPPSAFEMSGSDHNGSSASXISODVXKDEDFGSMKINIERGCTLTGELHAMELEKG      | 291   |
| Db | 1285 | VPPSPSPVFSFSSMGSDCAQAPATAVSDEDEKEDDFGSMKNIDERYGSLTGLQHLVLELKG   | 13444 |
| Qy | 292  | HSGLGLSLAGKXKDSRMSASVTVGIDPNGAAGKGRQIADDELLEINQOILTYRSRQNMSS    | 351   |
| Db | 1345 | QSGIGLSLAGKXKDSRMSASVTVGIDPNGAAGRGRQIADDELLEINQOILTYRSRQNMSS    | 14040 |
| Qy | 352  | IIKCAPSKVKIIFIRNKDAVNOMAVCPGNAVEPLPSNSENLONKEFPTVTTSDAAVDLS     | 411   |
| Db | 1405 | IIKCAPSKVKIIFIRNKDAVNOMAVCPGIAADSPSSSTSDSPONKEVEPCSTTSASADLS    | 14644 |
| Qy | 412  | SFKVNOHELPHKQOOGGIGIAISEEDPLSGYIISLSTLEHVAATDGRKXGDDIILAVDEI    | 471   |
| Db | 1465 | SLTVYOLLELPKDDGGIGIACEEDYINGMIESTLEHGAAKDKRLKRGDHIILAVDEV       | 15244 |
| Qy | 472  | VWGPPIEFKIFSLTAKMTVKLTIRHNPQSAVPSAAGAASEGKNSQSOLMPQOSP          | 531   |
| Db | 1525 | VAGCPVEKFISLKTAAATYKLTIVBRAENPACPAVPSATVYSGERKDSNQTPAV---AP     | 1581  |
| Qy | 532  | EPESIRNTSRSTPAIPASDPATCPIIPGCETTIEISKRTGLSLISVGSSTLLGAETI       | 591   |
| Db | 1582 | DLEPIPSRSTPAVPAASPATCPIIPGCETTIEISKQOTGLSLISVGSSTLLGATII        | 16411 |
| Qy | 592  | HEVEEBCAACCKDGLMLAGDDIILEVNGDILRKAHADAIVLQOTQORVRLTIRDEAYK      | 651   |
| Db | 1642 | HEVEEBCAACCKDGLMLAGDDIILEVNGDILRKAHADAIVLQOTQORVRLTIRDEAYK      | 17011 |
| Qy | 652  | EEBVCODTTLIELOKPKPGKGLSLIVGRKNDTVGFVEDIKYGSIADPDGRLIOGDDILLVN   | 711   |
| Db | 1702 | EBDVCODTTLIELOKPKPGKGLSLIVGRKNDTVGFVEDIKYGSIADPDGRLIOGDDILLVN   | 17611 |
| Qy | 712  | GEDYRNASQEAVALLKCSIGTYVTVLEYGRITAKPHESERRPSQTSOVSSEGSLSSTFFPLS  | 771   |
| Db | 1762 | GEDYRNHATQEAVALLKCSIGATVTVLEYGRYKAPAFHSERRPSOSSOVSSESSLSSTFFPLS | 18211 |
| Qy | 772  | GSSSTSESSLESSKKNALASEIOGIRTYEMKKGPTSDLSGISINGVGSPLGADVPITAMH    | 831   |
| Db | 1822 | GINTSESSLESSKKNALASEIOGIRTYEIKKGADSLGISINGVGSPLGADVPITAMH       | 18811 |
| Qy | 832  | PTGVAAGQTKLRVODRIVTTCSTSEBGMTRTQAVNLKLNASGSIEMQVAVAGDVSVMYTH    | 891   |
| Db | 1882 | PNGVAAGQTKLRVODRIVTTCSTSTOCSMTHTQAVNLKLNASGSIEMQVAVAGDVSVMYTH   | 19411 |
| Qy | 892  | HOEPASSLSFTGLTSTISIFODDLGPPOCKSITLERGPDGLGFTIVYGSGSPHGDPLIV     | 951   |
| Db | 1942 | QOELANPCLAFETGLTSSISIFPDGLGPPOCKSTITLDRGPDGLSPENIYVGSGSPHGDPLIV | 20011 |
| Qy | 952  | KTVYPAKGAASDGLRKGDQIIVANGSLGCVYHEEVAVALIKRTKCTVTLMLTS           | 1005  |
| Db | 2002 | KTVYPAKGAASDGLRKGDQIIVANGSLGCVYHEEVAVALIKRTKCTVTLMLTS           | 2055  |

| RESULT | 4   |              |      |         |
|--------|---|--------------|------|---------|
| 008783 |   |              |      |         |
| ID     | 008783  | PRELIMINARY; | PRT; | 526 AA. |
| AC     | 008783;   |              |      |         |
| DT     | 01-JUL-1997 (TREMBLrel. 04, Created)                              |              |      |         |
| DT     | 01-JUL-1997 (TREMBLrel. 04, Last sequence update)                 |              |      |         |
| DT     | 01-MAR-2001 (TREMBLrel. 16, Last annotation update)               |              |      |         |
| DE     | 90RF BINDING PROTEIN 1 (FRAGMENT).                                |              |      |         |
| GN     | 9BP-1.  |              |      |         |
| OS     | Mus musculus (Mouse).   |              |      |         |
| OC     | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; |              |      |         |
| OC     | Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus. |              |      |         |
| OX     | NCBI_TaxID=10090;   |              |      |         |
| RN     | [1]   |              |      |         |
| RP     | SEQUENCE FROM N.A.  |              |      |         |
| RC     | TISSUE=PANCREAS;  |              |      |         |
| RA     | Lee S.S., Weiss R.S., Javlier R.T.;                               |              |      |         |
| RL     | Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.           |              |      |         |
| DR     | EMBL; AF000168; AAB57835.1; -                                     |              |      |         |

DR HSSP: Q12959; 1PDR.  
 DR InterPro: IPR001478; -  
 DR Pfam: PF00595; PDZ; 4.  
 DR SMART: SM00228; PDZ; 1.  
 FT NON\_TER 1  
 SQ SEQUENCE 526 AA; 54948 MW; 3843CE3AA2410E8 CRC64;

Query Match 45.2%; Score 2297.5; DB 11; Length 526;  
 Best Local Similarity 86.6%; Pred. No. 5.6e-11;  
 Matches 457; Conservative 34; Mismatches 32; Indels 5; Gaps 2;

QY 480 FLSLKAKMTVKLTTHAENDSOAVPSAAGASGEKNSOSILMVPDGSPEBESINT 539  
 DB 2 FSLTKAKATVKLTTHAENDSOAVPSAAGASGEKNSOSILMVPDGSPEBESINT 58  
 QY 540 SRSSTPAIFASDPATCPPIPCETTIEISKRTGLSLVSGSDTLGAFIIHEYEEGA 599  
 DB 59 SRSSTPAIFASDPATCPPIPCETTIEISKRTGLSLVSGSDTLGAFIIHEYEEGA 118  
 QY 600 ACKGRLMAGDQILEVNGIDLRKATHDEAINVLRQTPORVRLTYRDEAPYKEEVCDTL 659  
 DB 119 ACKGRLMAGDQILEVNGIDLRKATHDEAINVLRQTPORVRLTYRDEAPYKEEVCDTL 178  
 QY 660 TIE--LQKPKGGLSLVGRNDTVFVSDIVKGIADPPDRLIOGDQILLNGEDVRN 717  
 DB 179 TIEHOLKRPKGLSLVGRNDTVFVSDIVKGIADPPDRLIOGDQILLNGEDVRN 238  
 QY 718 ASQAVALLKCSIGVYLVLEGRKAGPHERPSQTSQVSEGLSFTPLSGSSSE 777  
 DB 239 ASQAVALLKCSIGVYLVLEGRKAGPHERPSQTSQVSEGLSFTPLSGSSSE 298  
 QY 778 SLESSKKNALASEIQLRTVEEMKGPPTSIGISLIVAGVSPGLDVPFIAMHPGTGVA 837  
 DB 299 SLESSKKNALASEIQLRTVEEMKGPPTSIGISLIVAGVSPGLDVPFIAMHPGTGVA 358  
 QY 838 QTKRLVGRDRIYTCIGSTEGMTHTOAVNLKNAAGSIEMOVAVAGDVSVTGHQEPAS 897  
 DB 359 QTKRLVGRDRIYTCIGSTEGMTHTOAVNLKNAAGSIEMOVAVAGDVSVTGHQEPAS 418  
 QY 898 SLSFTGLTSTIFODDGPPOCKSITLERGPDGLGFSIVGSGSPHGLDIYKTVFAK 957  
 DB 419 PCLMFTGLTSSIFPDGPPQSKITLDRGPDGLGFSIVGSGSPHGLDIYKTVFAK 478  
 QY 958 GAASEDGRLKRGDQIIAVNGQSLGVTHEEAVAILKRTKGTVTLMVLS 1005  
 DB 479 GAASEDGRLKRGDQIIAVNGQSLGVTHEEAVAILKRTKGTVTLMVLS 526

RESULT 5  
 ID 043798 PRELIMINARY; PRT; 453 AA.  
 AC 043798;  
 DT 01-JUN-1998 (Tremblrel. 06, Created)  
 DT 01-JUN-1998 (Tremblrel. 06, Last sequence update)  
 DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)  
 DE MUPPI PDZ DOMAIN PROTEIN 1 (FRAGMENT).  
 GN MUPPI.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN 11  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-BRAIN.  
 RX MEDLINE=98196865; PubMed=9537516;  
 RA Ulmer C., Schumack K., Flege A., Lubbert H.;  
 RT "Cloning and characterization of MUPPI, a novel PDZ domain protein.";  
 RL FEBS Lett. 424:63-68(1998).  
 DR EMBL: AJ001319; CA004680.1; -  
 DR HSSP: Q12959; 1PDR.  
 DR InterPro: IPR001478; -  
 DR Pfam: PF00595; PDZ; 4.

DR SMART: SM00228; PDZ; 1.  
 FT NON\_TER 1  
 SQ SEQUENCE 453 AA; 47423 MW; DA7915AD493C5D8 CRC64;

Query Match 44.3%; Score 2252; DB 4; Length 453;  
 Best Local Similarity 98.2%; Pred. No. 2.5e-128;  
 Matches 445; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 553 ATCPPIPCETTIEISKRTGLSLVSGSDTLGAFIIHEYEEGAACKGRIMADQI 612  
 DB 1 ATCPPIPCETTIEISKRTGLSLVSGSDTLGAFIIHEYEEGAACKGRIMADQI 60  
 QY 613 LEVNGIDLRKATHDEAINVLRQTPORVRLTYRDEAPYKEEVCDTLIELQKPKGGLG 672  
 DB 61 LEVNGIDLRKATHDEAINVLRQTPORVRLTYRDEAPYKEEVCDTLIELQKPKGGLG 120  
 QY 673 LSIYGRNDTVFVSDIVKGIADPPDRLIOGDQILLNGEDVRNASEAVALLKCSIG 732  
 DB 121 LSIYGRNDTVFVSDIVKGIADPPDRLIOGDQILLNGEDVRNASEAVALLKCSIG 180  
 QY 733 TVTLEVGRIKAGPHERPSQTSQVSEGLSFTPLSGSSSELESSKKNALASEI 792  
 DB 181 TVTLEVGRIKAGPHERPSQTSQVSEGLSFTPLSGSSSELESSKKNALASEI 240  
 QY 793 QGLRTVEEMKGPPTSIGISLIVAGVSPGLDVPFIAMHPGTGVAQTKLVRGRIYTC 852  
 DB 241 QGLRTVEEMKGPPTSIGISLIVAGVSPGLDVPFIAMHPGTGVAQTKLVRGRIYTC 300  
 QY 853 GSTEGMTHTOAVNLKNAAGSIEMOVAVAGDVSVTGHQEPASSLSFTGLTSSIFQ 912  
 DB 301 GSTEGMTHTOAVNLKNAAGSIEMOVAVAGDVSVTGHQEPASSLSFTGLTSSIFQ 360  
 QY 913 DDLGPPCKSTTLERGPDLGFSIVGSGSPHGLDIYKTVFAKGAASEDGRLKRGDQI 972  
 DB 361 DDLGPPCKSTTLERGPDLGFSIVGSGSPHGLDIYKTVFAKGAASEDGRLKRGDQI 420  
 QY 973 IAVNGQSLGVTHEEAVAILKRTKGTVTLMVLS 1005  
 DB 421 IAVNGQSLGVTHEEAVAILKRTKGTVTLMVLS 453

RESULT 6  
 ID 09H3N9 PRELIMINARY; PRT; 1134 AA.  
 AC 09H3N9;  
 DT 01-MAR-2001 (Tremblrel. 16, Created)  
 DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)  
 DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)  
 DE PDZ DOMAIN PROTEIN 3' VARIANT 4.  
 GN HINADL.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN 11  
 RP SEQUENCE FROM N.A.  
 RA Soejima H., Kawamoto S., Akai J., Miyoshi O., Arai Y., Nikiawa N.,  
 RA Kimura A., Okubo K., Mukai T.;  
 RT "Cloning, characterization and chromosomal mapping of the two novel  
 RT heart specific genes.";  
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AB044807; BAB19683.1; -  
 SQ SEQUENCE 1134 AA; 125244 MW; FB6C3DA92D7CEFB6 CRC64;

Query Match 29.4%; Score 1496.5; DB 4; Length 1134;  
 Best Local Similarity 44.3%; Pred. No. 5.3e-82;  
 Matches 360; Conservative 99; Mismatches 177; Indels 177; Gaps 21;

QY 3 ONVSKESPER-----TINAKGNSLGMVTSANKKGLGMIYSHHGAIASDGRITAGDC 58  
 DB 402 ENVKKNFVMSLPSVSTSGNSQGR--FDDLENLSLAKTSLDGLKIPND----- 451



```

QY 59 ILINBESTISVTNQAARMLRHSLLIGPDIKITY-VPAHLEEKISIG--QSGRYMAL 116
DB 452 -----VOGSLIIDLPPVAAQREOEDLPYHOQARRVISK 486
QY 117 DISSVTG-----RDIPERECEGESELONTATASNNOPRVLMEEPSKSLGI 169
DB 487 --ASAVTGLSSRYATDICELEPEREGEGER-----PNFSHMGPRIVIEIFREPNVSLGI 540
QY 170 STVGGRGMSRLSNGEVMGIFIKHYLEDSPAGKNGTLPGRIVFA-----PSOSESEP 224
DB 541 STVGQTVIKRLKNGELGIFIKVLEDSPAKTNALTKGILLEVSGVDLQNMASHSEA 600
QY 225 EKA-----PLCSVPPPPSAFAENG-----SDHTQSSASK----- 254
DB 601 VAIKNAGPVPVFIYOSLSTPRVLPVHNKANKITGONODOTKEKKEKROGTAPPKML 660
QY 255 -----ISODVDR-EDEFYSWKNIERYGTGLGELHMLEKSHGSLGLSLAGNKRGR 307
DB 661 PPYKALTDSDSENEEDAFDTQKINQRYADLPGLHITILEKDKNGLSLAGNKRGR 720
QY 308 MSVFIYGDIPNGACGKDLQIADLELTNGOILYGRSHQNASITKCAPSKVKITIFIRN 367
DB 721 MSIFVYGINPEGRPAADGEMHIGDELLEINQOILYGRSHQNASITKCAPSKVKITIFIRN 780
QY 368 KQAVNOMAVCPGNAVEPLSENENONKTEPEPTVTSDAAVDLSSFKNVQHLLEPKDGG 427
DB 781 EAVNOMAVTP-----FPVSSSP-----SST-----EDQSG 807
QY 428 LGIAISEEDTLGVLIIKSLTEHGVATDGRKYDQILAVDEIYVGYPIEKFISILKTA 487
DB 808 TERPISSE-----DLSLEVGIKQLPSESFKLA-----VSGMKQO 842
QY 488 KMTVKLTTHAENPDQAVPSAAGASGEEKSSOSLMPVQSSPEPESIRNTSRSTPAI 547
DB 843 KPTKXSF-----SSQEIPLAPASS-----YHSTDADFTGYGFOAPLVS----- 882
QY 548 FASDPATCPTIIGCETTITISKRGGLGSLYSGSPTLLGAFIHYVEEGAACKRGRLW 607
DB 883 ---DPTCPIVPGQEMIILISIKRSGSLSYGVGKDTPLNATVILHYVEEGAARGRGRL 939
QY 608 AGDQILEVNGIDLKRTKDEALINVLROTPOVRLTYRDEAPYKEEVEDDTLTIELQKRP 667
DB 940 AGDQILEVNGIDLKRNSSHEEATLALROTPOKRLVYRDEAHYRDEENLEIFPVDLQKRA 999
QY 668 GGLGLSLYGVKRNQDGVFVSYIVKGIADPDGLIQQDQILLVNGEDVNASQEAVALLL 727
DB 1000 GGLGLSLYGVKRNQDGVFVSYIVKGIADPDGLIQQDQILLVNGEDVNASQEAVALLL 1059
QY 728 KCSLGTVITLEVGRIRKAPFHSRRPSQTSQVSE 760
DB 1060 KCAOGLVQLEIGRLRAGSWTSATTSONSOAE 1092

```

```

RESULT 7
070471 PRELIMINARY: PRT: 612 AA.
AC 070471:
DB 01-AUG-1998 (TREMBLrel. 07, Created)
DB 01-AUG-1998 (TREMBLrel. 07, last sequence update)
DB 01-MAR-2001 (TREMBLrel. 16, last annotation update)
DB CHANNEL-INTERACTING PDZ DOMAIN PROTEIN.
GN CIPP.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BALB/C; T1SSUE-CEREBELLUM, BRAIN;
RA Kurchner C., Mermelstein P.G., Holden W.T., Surmeier D.J.;
RL Mol. Cell. Neurosci. 0:0-0(1998).
EMBL, AF060539; AAC40148.1; -.

```

```

DR HSSP: 012959; 1PDR.
DR MGD; MGI:1277960; CAPP.
DR InterPro; IPR001478; -.
DR Pfam; PF00595; PDZ; 4.
DR SMART; SM00228; PDZ; 1.
SQ SEQUENCE 612 AA; 65420 MW; 145DE1769F54BE56 CRC64;

Query Match 28.9%; Score 1469.5; DB 11; Length 612;
Best Local Similarity 49.5%; Pred. No. 8.5e-81;
Matches 335; Conservative 85; Mismatches 158; Indels 99; Gaps 12;

QY 270 KNIRERYGTLTGELHMLEKSHGSLGLSLAGNKRGRSVTVIGIDPNGACGKRLQI 329
DB 8 EKTORRYADLPGLHITILEKDKNGLSLAGNKRGRSVTVIGIDPNGACGKRLQI 67
QY 330 ADELLEINQOILYGRSHQNASITKCAPSKVKITIFIRNKDAVNOMAVCPGNAVEPLPSNS 389
DB 68 GDELEINQOILYGRSHQNASITKCAPSKVKITIFIRNKDAVNOMAVCPGNAVEPLPSNS 116
QY 390 ENLONKTEPEPTVTSDAAVDLSSFKNVQHLLEP-KDQGLGIAISEEDTLGVIILKSLTE 448
DB 117 -----FPELSHSPSPVEDLGCTELVSSSE--SSVDAKHLPE 151
QY 449 HGVAAATDGRKYDQILAVDEIYVGYPIEKFISILKTAKMVTKLTHAENPDQAVPSA 508
DB 152 PSS-----KEDLSQVYVDNMVAEOKE-----SESPDSACQIK 187
QY 509 AGASGEEKKSSOSLMPVQSSPEPESIRNTSRSS-----TPAIRSPADPTPIIP 559
DB 188 QQTYSQVSSSQD--SPSPAPLQGSAAHADVTGSGNFQAPLPVDPAPLSPDPTPIIP 245
QY 560 GCETTITISKRGGLGSLYSGSPTLLGAFIHYVEEGAACKGRLMAGDQILEVNGID 619
DB 246 GQEMIIEISKGRSGGLSLYGVGKDPPLDAIVHYVEEGAARADGRLMAGDQILEVNGID 305
QY 620 LKRTKDEALINVLROTPOVRLTYRDEAPYKEEVEDDTLTIELQKRPKGLSLYGR 679
DB 306 LRSSSHEEATLALROTPOKRLVYRDEAHYRDEENLEIFVLDQKTKRGGLSLYGR 365
QY 680 NDTGVFVSDIVAGIADPDGRLIQQDQILLVNGEDVNASQEAVALLLKCSLGTVITLEV 739
DB 366 SSGSVFISDIVGAADLDGRLIRGQDILSVNGEDMRHNSQETVATILKCYGVQVLEIG 425
QY 740 RIKAGPFHSRRPSQTSQVSEGLSFTPLPSGSTSELES-----SSKNALASEIQ 793
DB 426 RLRAGSWASRKTSONSQDQSHASSCRP--SFAPVTSIQMLVYGRKSSDPPOKTE-E 483
QY 794 GLRTVEMKGPDSIGSLAGVSPLDGVPPIFIAMHPTGYAAGTOKLRVGDRTVITIG 853
DB 484 EPTVEELIELSDALGSLTAGKSPLDGPIPIFIAMIQANGVAARKOKLKVGDRTVITIG 543
QY 854 TSTEGTHTQAVNLKLNAGSIEMQVAGADVVTYGHQHPASSSTFTGLTSTIFED 913
DB 544 QPLDGLSHTDANVLKNAGGRILLQVYVADTNISALATOLEINSAGS----- 589
QY 914 DLGPPQCKSITLERGPD 930
DB 590 QLGSP-----TADRHE 601

```

```

RESULT 8
060833 PRELIMINARY: PRT: 1552 AA.
AC 060833:
DB 01-AUG-1998 (TREMBLrel. 07, Created)
DB 01-AUG-1998 (TREMBLrel. 07, last sequence update)
DB 01-MAR-2001 (TREMBLrel. 16, last annotation update)
DB INADL, C-TERM VARIANT2.
GN INADL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

```



```

Db 1349 TEPISSE-----DGSLEVGIKQLPESSEFKLA-----VSQMKQ 1383
Oy 488 KMTVKLTTHAENPDQAVPSAAGASGEKKNSSOSLWVPOSGSPPEISIRNTSRSTPAI 547
Db 1384 KYPTKVSF-----SSQEIPLAPASS-----YHSTDADEFTGYGFOAPLSV-----1423
Oy 548 FASDPATPIIPGCTTETIISKGRGLSTVGSDDL-----586
Db 1424 ---DEATCPPIYVGQEMITETISKGRGLSTVGSKDTPLFWRLGSPRAMSOHLVRAFMILH 1480
Oy 587 -----GAFIIHEVEEGAACDKGRMLAGDQILEVNGIDLRKATHDEATINVRQTPQ 637
Db 1481 HPVTEGONALVIHVEYEEGAARDGRMLAGDQILEVNGVDLRNRSHEATLALROTPO 1540
Oy 638 RVRLLTYRDEAPYKEEYCDTLTIELOKKPGKGLSTVGR 679
Db 1541 KYRLVYRDEAHYDEENLEIFPVLDQKKAGRLGLSTVGR 1582

RESULT 10
ID 015249 PRELIMINARY; PRT; 1524 AA.
AC 015249;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE PDZ DOMAIN PROTEIN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
NCBI_TaxID=9606;
RN [1]
RP MEDLINE=97423468; PubMed=9280290;
RA Philipp S., Flockerzi V.;
RT "Molecular characterization of a novel human PDZ domain protein with
RT homology to INAD from Drosophila melanogaster.";
RL FEBS Lett. 413:243-248(1997).
RN [2]
RP MEDLINE=96224170; PubMed=8617505;
RA Lennon G., Aufray C., Polymeropoulos M., Soares M.B.;
RT "The I. M. A. G. Consortium: an integrated molecular analysis of
RT genomes and their expression.";
RL Genomics 33:151-152(1996).
DR EMBL; AJ001306; CAA04666.1; .
DR HSSP; P31016; 1BFE.
DR InterPro; IPR001478; .
DR Pfam; PF00595; PDZ; 8.
DR SMART; SM00228; PDZ; 1.
SQ SEQUENCE 1524 AA; 167323 MW; E3BBA7EC856A954D CRC64;

Query Match 20.9%; Score 1063.5; DB 4; Length 1524;
Best Local Similarity 38.4%; Pred. No. 1.3e-55;
Matches 281; Conservative 87; Mismatches 155; Indels 205; Gaps 22;

Oy 3 QNVSSESTER---TINIAKNSLSGMTVSAKNDGLMIVRSIIHGALSRDGRITAGDC 58
Db 943 ENVMKENFVMSLPSVPSTEGNSQGR--FDDLEMLNSIAKSLDLMIPND-----992
Oy 59 ILSTMEESTISTYNAQARMLRRLSLIGPDIKITY-VPAEHLEERKISLQ-QSGRWAL 116
Db 993 -----VQGSLLIDLPVVAQREQDLPYQHQAATRVISK 1027
Oy 117 DISSSYTG-----RDIPELPEREGEGESELQNTAVSNMNPRAVELMREPSKSLGI 169
Db 1028 ---ASAVTGLMSLRVATDICELPREGEGER---PNFSHMGPRIVIEFREPVSIGI 1081
Oy 170 STVGRGMSRLSNCEVNMKGIFIKHVLSDPAKNGTKLPGGRIVEA-----PSQSESEP 224
Db 1082 STVGQATYKRLKNGEELKGIKIFKOVLEDSPAKTNALKTGDKILEVSGVDLQNSHSPA 1141

```

```

Oy 225 EKA-----PLCSVPPPPSAFAEMG-----SDHTOSSAK-----254
Db 1142 VEAIRNAGNPVYFIVQSLSSTPRVTPNVNANKITGQNONDQTKKERRQGTAPPPKL 1201
Oy 255 -----ISODVDR-BDEGYSWKINRERYGTLTGLHIELEKSGSLGSLAGKDKSR 307
Db 1202 PPPYALRDDSDENEEDAFDQKTRQRYADLPGLHITIELEKNGSLGSLAGKDKSR 1261
Oy 308 MSVFLVIGIDIPNGAACKDGRLOJADELLEINQOILGRSHQNASIIKCAPSKVITIFRN 367
Db 1262 MSIFVGVINPEGRPAADGRMHIGDELLEINQOILGRSHQNASIIKCAPSKVITIFRN 1321
Oy 368 KDVAVMQAVCGNAVEPLPSNSENIQNKETEPTVTSAAVDLSFKVVOHLEPKDGG 427
Db 1322 EDVAVMQAVTP---FPVSSSP-----SSI-----EDSG 1348
Oy 428 LGIAISEDTLSGVITIKSLTEHGVATDGRGLKVGDOILAYDDEIVGVPIEKFISLTKTA 487
Db 1349 TEPISSE-----DGSLEVGIKQLPESSEFKLA-----VSQMKQ 1383
Oy 488 KMTVKLTTHAENPDQAVPSAAGASGEKKNSSOSLWVPOSGSPPEISIRNTSRSTPAI 547
Db 1384 KYPTKVSF-----SSQEIPLAPASS-----YHSTDADEFTGYGFOAPLSV-----1423
Oy 548 FASDPATPIIPGCTTETIISKGRGLSTVGSDDLGAFTIHEVEEGAACDKGRML 607
Db 1424 ---DEATCPPIYVGQEMITETISKGRGLSTVGSKDTPL-----1459
Oy 608 AGDQILEVNGIDLRKATHDEATINVRQTPORVRLTYRDEAPYKEEYCDTLTIELOKK 667
Db 1460 -----VNGVDLRNRSHEATLALROTPOKYRLVYRDEAHYDEENLEIFPVLDOKKA 1512
Oy 668 GKGGLSTVGR 679
Db 1513 GRGLSTVGR 1524

RESULT 11
ID 009515 PRELIMINARY; PRT; 2208 AA.
AC 009515;
DT 01-JUN-1998 (TREMBLrel. 06, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE HYPOTHEICAL 240.1 KDA PROTEIN C52A11.4 IN CHROMOSOME II.
GN C52A11.4.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabdilita; Rhabdilitidae;
OC Rhabdilitidae; Peloderinae; Caenorhabditis.
NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL NZ;
RA Coles L., Sulston J.;
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; Z46792; CAA86769.2; JOINED.
DR EMBL; Z46795; CAA86769.2; JOINED.
DR EMBL; Z46792; CAA86789.2; JOINED.
DR HSSP; Q12959; 1PBR.
DR InterPro; IPR001478; .
DR InterPro; IPR002173; .
DR Pfam; PF00595; PDZ; 15.
DR PROSITE; PS00584; PKB_KINASES_2; UNKNOWN.1.
SQ SEQUENCE 2208 AA; 240120 MW; B6894F8D603E79FB CRC64;

Query Match 16.0%; Score 814; DB 5; Length 2208;
Best Local Similarity 20.9%; Pred. No. 2.9e-40;
Matches 307; Conservative 137; Mismatches 259; Indels 768; Gaps 33;

Oy 11 ERTINIAKNSLSGMTVSAKND-GL-GMIVRSIIHGALSRDGRITAGDCILSTINEESTI 68

```

```

Db 895 ERTYKLGALPLGAVLDGDKDKNVCVVKSIGCKAKVALDRIQVGDFTIKINTESLR 954
QY 69 SUTNAOARMLRRSLIGPDKITVYP----- 95
Db 955 NVTNSOAAALIKRLNLTGTCNVTYITTSADAKTKERNRORPSESSPIINLSKRVPEKE 1014
QY 96 -----AHLBEFKISLG-----OOSGRVM 114
Db 1015 YRSPFMQROESQSKTEMTDDETAAPSIMTDSMSHIKTFDLAEGSSNSHHDEQVNRMS 1074
QY 115 AL-----DIF 119
Db 1075 RLIDGVEVDVNLILKEAITDATIELRLVHLKTKDWSNCKNRERLESPLPPPEVL 1134
QY 120 SS-----YTGRDI-----PELPE 132
Db 1135 SSPKSPVAVQPTPRDLEEVLTSTASSLEFHSQRTSQLHLSTEBEVLQATPPSSPE 1194
QY 133 REEGE----- 137
Db 1195 NKSEVEPSPISPGIKLAGEVTAPBEIEVVEQAENDRAETANGAEVETSTPTAEAI 1254
QY 138 -----GESELCNTA-----Y 148
Db 1255 GNSKDESTTTSISQOSVGLQTOALNSTEVNSNMSRYTSRTPTSGSES-LQNOAROLVRS 1313
QY 149 SNMNOPRVELMREPSKSLGISVIGR-----GMGSRLSNGEVMGIFIKHVLDEDSAPAKN 204
Db 1314 KYWGEARTVTLVREPNKSFSGISIVGRVEVQKGLPETGNTVCGIFIKSVLPNSPAKRS 1373
QY 205 GTLPGDRIVEAP----- 217
Db 1374 GOMNMDRIVISVNDVLDLADTAHQAVNAIKMNSNPVFLVLSLHTNOOMINSASNTVG 1433
QY 218 -----SSESEPEKAPLCSVPP-----PPPS-----AFKMSDHTQS 250
Db 1434 SVRENNKPEEELPPTALVPLKPMISSGSGSTSKPAMNPPPSISTTTTSMESKEE 1493
QY 251 SAKRISODVKED----- 263
Db 1494 EATSSSPDIQENVTYKRRSMOVERQEPANVEETSIYEKTPPRKTSAKNSQKESND 1533
QY 264 ----- 263
Db 1554 RKEIKKQKVEREMSVESKKSRSIKKHROESTIIKKSNETAPLYSDVSSETHDE 1613
QY 264 -----EFGYSKWNIRRYTGLGEL 283
Db 1614 POAMSPSTSPDTRDADAMRALGIDDSAAFOIKNDGEPSKFFYTAGIERKYPDSGEL 1673
QY 284 HMLEKGHSGLSGLAGNKRSMASVPIYIGIDPNGAAGKGRLOIADELEILGOLYG 343
Db 1674 VLVACERPDGGLGSLGNKRDKNQVFNVRSCPLA-----TRPGELLEILGRLINK 1729
QY 344 RSHONASSIIR-CAP--SKVYIFIRKKDAVNMQAVCPGNAVEBPLPSNSENLOKRETEPT 400
Db 1730 ISHVAASAVARECCDQONIEITVLRNRGALKSA-----PQMPTTBEIEN----- 1776
QY 401 VTTSDAAVDLSSFKNVQHLPLKQDGLGIAISEBTLGVIYIKSLTEHGVAAATDRLKV 460
Db 1777 -----AAPNOEELS----- 1785
QY 461 GDQILAVDELIVGPIEKFISLTKTAKMYKLTIAHBNPDSQAVPSAAGAASEKKNS 520
Db 1786 -----RKKSFS 1791
QY 521 QSLWPOSGSPESIRNTSSSTPAIFASDPATPIPGCETIIEISKRTGLSLG 580
Db 1792 QE-----RTQATEN-----GRETMIELIDKQKGLGSLG 1821
QY 581 GSDTLGAFIIEHYEGGAACKGRRLNAGDOILEVNGIDLRKATHDEALNVLROTPOVR 640

```

```

Db 1822 GADTVLGTVVIHEEYSDGAAAHDRKLPQDQVLEVNQSTSLRGVTHDSIAYLRRTPPKVR 1881
QY 641 LTLRYDE-----APYKEEVCOTLTIELOKKRGKGLGISTYCKRNDGVFVSDIYKGIAD 696
Db 1882 LLITRDVNLQSLDLPQIYIIFEIDLKTKGRLGSLISYGRKNEPQVYSEIYKGLAE 1941
QY 697 PDGRLLIOGDOILLVNGEDVNRASQEAVALLKSLGTVTLEVGRIKAPF----- 746
Db 1942 SDGRIMTGDQLLEVNGKDVRCMQEDVAAMLKTTGKVLKTEBNNNDPVPAVAVSSP 2001
QY 747 ----- 746
Db 2002 AATPSTASVTVTSAPATSVPEPNVVEPIPGPTSAARENDVPPAPPMKPIITHTSPE 2061
QY 747 --HSERRPSOTSOVSEGLSFTF-----PLSGSTSELSSSKKNALASEI----- 792
Db 2062 GCETQOEPAGSLPYTERPSSCDPMVOEEOQPTPTTTT-----SSNNNSLADIIHDL 2118
QY 793 --OGLRT--VEKKGPDSLSGISIAGVSPLDGVPFIAMHPTGVAATOQLRVGDRI 848
Db 2119 KEESDQLVELKVKVDOQLGM-----GIGK--RSRGILVTSLOPSSAA--EKLKVDRI 2170
QY 849 VTICSTEGMTHQAVNLKKNAGSIEMQY 879
Db 2171 LAV--NALPVSQLSAVTEVAKASGORLYLOI 2199

RESULT 12
070263
ID 070263 PRELIMINARY; PRT; 728 AA.
AC 070263;
DT 01-AUG-1998 (Tremblrel, 07, Created)
DT 01-AUG-1998 (Tremblrel, 07, Last sequence update)
DE 01-MAR-2001 (Tremblrel, 16, Last annotation update)
DE NIMB-BINDING PROTEIN LNX (LIGAND OF NIMB-PROTEIN X).
GN LNX.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A. AND CHARACTERIZATION.
RC TISSUE-BRAIN, AND EMBRYO;
RX MEDLINE=98204916; Pubmed=9535908;
RA Dho S.E., Jacob S., Wolfing C.D., French M.B., Rohrschneider L.R.,
RA McGlade C.J.;
RT "The mammalian numb phosphotyrosine-binding domain. Characterization
RT of binding specificity and identification of a novel PDZ domain-
RT containing numb binding protein, LNX."
RL J. Biol. Chem. 273:9179-9187(1998).
CC -FUNCTION: INTERACTS WITH THE PHOSPHOTYROSINE INTERACTION DOMAIN OF
CC THE NIMB PROTEIN.
CC -ALTERNATIVE PRODUCTS: TWO ISOFORMS; ISOFORM LNX (OR LNXP80) (SHOWN
CC HERE) AND ISOFORM LNX-B (OR LNXP70); ARE PRODUCED BY ALTERNATIVE
CC SPLICING.
CC -TISSUE SPECIFICITY: BOTH ISOFORMS ARE EXPRESSED IN THE HEART.
CC ISOFORM LNX IS ALSO EXPRESSED IN KIDNEY, LUNG AND SKELETAL MUSCLE
CC WHILE ISOFORM LNX-B IS ALSO EXPRESSED IN BRAIN.
CC -SIMILARITY: CONTAINS 4 PDZ DOMAINS.
EMBL: AF034745; AAC40075.1; -
DR EMBL; AF034746; AAC40076.1; -
DR HSSP; P31016; 1BFE.
MGI: 1278335; Lnx.
DR InterPro: IPR001478; -
DR Pfam; PF00097; zf-C3HC4; 1.
DR SMART; SM00228; PDZ; 1.
DR Pfam; PF00595; PDZ; 4.
DR SMART; PF00595; PDZ; 4.
DR Alternative splicing.
FT DOMAIN 278 362 PDZ.
FT DOMAIN 385 467 PDZ.
FT DOMAIN 508 593 PDZ.
FT DOMAIN 638 723 PDZ.
FT MOTAGEN 188 188 Y->A: ABOLISHES BINDING TO NIMB PROTEIN.

```

FT MUTAGEN 188 188 Y->F: NO EFFECT ON BINDING TO NUMB  
 FT VANSPLIC 1 131 PROTEIN  
 FT 130 TSCKASHGTLTKDRKRSQDCPDGCASLMTLIS-----PEVSAAT----- 173  
 FT 464 ILAVDEIVGYPIEFKELITAKMTVLTITHAENPOQAVP---SAGAASGEKKNS 520  
 FT 174 ISLMTDEPGLNP--AYSVSEDEGEVANSSDGSKNRTARPERSTKRSFKINRA 231  
 FT 521 QS-LMWPOSGSPESPESI---NTSKRSSTPAIFASDPATCPIPGCE--TTIEISKGRIGL 574  
 FT 232 LSLARTKSGSVYANHVDOGRDNSEHTVPEVF---PRLFHLIPDELTSIKINRADPSE 288  
 FT 575 GLST---VGSOTLLGAFITHEVEEGACKDRLNAGDQILVNGDILKKAHDEANLV 632  
 FT 289 SLSILVGGSEPLVHIITIIQITVROGVIAARDRLLPGLDILVNGDISNVPHNYAVRL 348  
 FT 633 RQTPORVRLTYLDE-----APYKEEVEGDTLITELQK-KPKGKILVIGKRRNT 682  
 FT 349 RQPCQVRLTYLRECKFKRSNAHPDSYGRDDSHVTLKNSPEQGLITLVRRVDEP 408  
 FT 683 GVFSVDIVKGIADPDGRLIQGDIILLVNGEDVNRNASQGAVALLKSLGVTLEVGRIG 742  
 FT 409 GVFIENVLNGVADRHQGLENDRLVLAINGHDLRFGSPSAHLIQASBRVHLVYSR-- 466  
 FT 743 AGRFHSERPSP-----QTQVSEGSLSSTFFPLSGSSTESLESSESKKNAALSEIOGLAT 797  
 FT 467 -----QVRQSPDIFQEAQWISNGOOS---PGPGR---NTASKPATCHE---KV 508  
 FT 798 VEMKAGPTSLGISIAGVSPGLDVPFIAMNHPTGVAQOTOKLEGRITVIGCTSTRE 857  
 FT 509 VSVKMDPSLSLGMTVGGGASHREMDLPITVIVSEPGVTSRGRITGIGILLANVCIETL 568  
 FT 858 GMTHTQAVNLKKNASGISIMOVYAGDVSVTGHQDEPASSLSLSTGLTSTIFODDLGP 917  
 FT 569 EVSRTAEVAILKSPSSVYLKAL---EVK-----EQEAQEDCSPALDS---NHNVT 615  
 FT 918 P-----QCKSTLERGPDG-LGFSIVGIGVSPRGDLPITYVKYFAAG 958  
 FT 616 PEDWSPSVMMLELPOLYLCNCKDVILIRNTASLGFICYGVEYSGNPFPIKTSIVEGT 675  
 FT 959 AASEGRLKRGDQIITAVNGQSLEGVTHEEAVAILKRTKGTIVLMVLS 1005  
 FT 676 PAVNGRIRCGDILLAVNGRSTSGMTHACLAKMKELKRITLTIAS 722

Query Match 11.6%; Score 589; DB 11; Length 728;  
 Best Local Similarity 29.1%; Pred. No. 2e-27;  
 Matches 188; Conservative 115; Mismatches 239; Indels 106; Gaps 23;

RESULT 13  
 ID 023823 PRELIMINARY; PRT: 665 AA.  
 AC 023823  
 DT 01-NOV-1996 (TREMELREL. 01, Created)  
 DT 01-NOV-1996 (TREMELREL. 01, Last sequence update)  
 DT 01-MAR-2001 (TREMELREL. 16, Last annotation update)  
 DE INAD PROTEIN.  
 GN INAD.  
 OS Calliphora vicina (Blue blowfly) (Calliphora erythrocephala).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Oestridae; Calliphoridae; Calliphora.  
 OC NCBI\_TaxID=7373;

RN [1]  
 RP SEQUENCE FROM N.A., AND CHARACTERIZATION.  
 RC STRAIN=CHALKY; TISSUE=RETINA;  
 RX MEDLINE=96216082; PubMed=8662634;  
 RA Huber A., Sander P., Paulsen R.;  
 RT "Phosphorylation of the Inad gene product, a photoreceptor membrane  
 FT protein required for recovery of visual excitation.";  
 RL J. Biol. Chem. 271:11710-11717(1996).  
 CC -I- FUNCTION: MAY BE INVOLVED IN CONTROL OF THE LIGHT RESPONSE. A RISE  
 CC IN INTRACELLULAR CALCIUM LEVELS UPON VISUAL EXCITATION MAY  
 CC INITIATE PHOSPHORYLATION OF THE INAD PROTEIN BY EYE-PKC.  
 CC PHOSPHORYLATED INAD MAY IN TURN ACT ON ANOTHER PROTEIN SUCH AS TRP  
 CC OR NORPA WHICH BOTH CO-PRECIPIRATE WITH INAD.  
 CC -I- SUBCELLULAR LOCATION: MEMBRANE-ASSOCIATED.  
 CC -I- TISSUE SPECIFICITY: EXPRESSED IN RHABDOMERAL PHOTORECEPTORS.  
 CC -I- SIMILARITY: TO DROSOPHILA INAD.  
 DR EMBL: 269889; CAA93758.1;  
 DR HSSP: P31016; 1BFE.  
 DR InterPro: IPR001478;  
 DR Pfam: PF00595; PDZ; 5.  
 DR SMART: SM00228; PDZ; 1.  
 KW Repeat, phosphorylation; glycoprotein.  
 FT DOMAIN 41 420 2 x APPROXIMATE REPEATS.  
 FT REPEAT 41 82 1.  
 FT REPEAT 381 420 2.  
 FT CARBOHYD 5 5 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 162 162 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 272 272 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 292 292 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 303 303 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT MOD\_RES 15 15 BY PKC (POTENTIAL).  
 FT MOD\_RES 168 168 BY PKC (POTENTIAL).  
 FT MOD\_RES 185 185 BY PKC (POTENTIAL).  
 FT MOD\_RES 320 320 BY PKC (POTENTIAL).  
 FT MOD\_RES 321 321 BY PKC (POTENTIAL).  
 FT MOD\_RES 437 437 BY PKC (POTENTIAL).  
 FT MOD\_RES 595 595 BY PKC (POTENTIAL).  
 FT MOD\_RES 658 658 BY PKC (POTENTIAL).  
 SQ SEQUENCE 665 AA; 73349 MW; C453192E1820ZBF7 CRC64;

Query Match 11.1%; Score 564.5; DB 5; Length 665;  
 Best Local Similarity 26.8%; Pred. No. 5.3e-26;  
 Matches 190; Conservative 109; Mismatches 229; Indels 181; Gaps 27;

QY 153 QPRVRLMREPSKSLGISIVGGRGMSRLNGEVNKGIFIKHVLDEDSPAKNGKTLKPGDR 212  
 DB 10 QVOSTYLDKTKGKSGSLIVRGEA-----RDSNSKGIYIKGIVPDSPLGLGKIKVGR 64  
 QY 213 IYEARS---QSESEP----- 225  
 DB 65 LTLNGKDVADTEPREVINLIKQAGSKIDLEQTYGSEQSNKGMEIKENGFSQNRNMEN 124  
 QY 226 ---KAPLCSVP-----PPPSAFAEKGSDHTOSSAS----- 253  
 DB 125 QDSINQPTKQRAIKRQSKMAQAPGRPVNNAKLNSTNTTSSKSDQDLDDDEDETRDM 184  
 QY 254 -----KISQVDK--EDEFYGMKNIIRRYGTLGELHMIEL 288  
 DB 185 TGRIRTAGVEIDRASAGNCKLNKMEKDRDKETDEDFGTAKINKRYVTMR-DLKKEI 243  
 QY 289 EK-GHSGGLSLAGKNDKRSMSFYIVGIDPNG-AAGKDRLOIADLELFINQIILYGRSH 346  
 DB 244 VPTNTALTLAAGSDRKMGCFAVGVNTSGPLASVD--IKSGEILLENSTVYLNKRGH 301  
 QY 347 ONASSIICAPSKVKIIFIRNKDAVNOMAVCPGNAVEPLPSNSENLQNKETPTVTTSDA 406  
 DB 302 LNASYIFKNIDE-RVLITSRRKRPND---EGMSVKP-----KKFPEI--DDT 345  
 QY 407 AVDLSEFKNVQHLLEPKDQGLGIAL--SEBDTISGVITKSLTEHGAATDRLKVGDO 463  
 DB 346 KFLFEQYAKARASVOY-KKEGFLGIMVYGVKHEVNGGIRITDLRESNMLAG-LKVGDM 403



RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Laeko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Mekulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusser D.R., Pacle J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Sidenklamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zheng G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of *Drosophila melanogaster*";  
 RL Science 287:2185-2195(2000).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-COMPOUND EYE;  
 RA Lindsley D.L., Zimm G.G.;  
 RL "The genome of *Drosophila melanogaster*";  
 RT submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-COMPOUND EYE;  
 RA MEDLINE-9512723; PubMed-7826638;  
 RA Shieh B.H., Niemeyer B.;  
 RT "A novel protein encoded by the *Inad* gene regulates recovery of visual  
 RL Neuron 14:201-210(1995).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-COMPOUND EYE;  
 RA Shieh B.;  
 RL submitted (OCT-1994) to the EMBL/GenBank/DBJ databases.  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-COMPOUND EYE;  
 RA Shieh B.;  
 RL submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AE003458; AAF46915.1; -;  
 DR EMBL: U15803; AAC36490.1; -;  
 DR HSSP: P31016; 1BFE  
 DR FlyBase; FBgn0001263; *Inad*.  
 DR InterPro; IPR001478; -;  
 DR Pfam; PF00595; PDZ; 5.  
 DR SMART; SM00228; PDZ; 1.  
 SO SEQUENCE 674 AA; 74332 MW; DAC24091D99EA7F3 CRC64;

Query Match 10.8%; Score 551.5; DB 5; Length 674;  
 Best Local Similarity 26.3%; Pred. No. 3,3e-25;  
 Matches 197; Conservative 133; Mismatches 285; Indels 135; Gaps 30;

OY 19 GNSSLGATV-----SANKDGLMTVRSIIHGCAISRDGRIAGDCLISNEESTISVT 71  
 DB 24 GKRSFGICIVRGVDSPTKTGTGIFIGIVDSPAHLGRLKGLKVDRLISLNGK--DVR 80  
 OY 72 NAQARAMLRRHSLIGPDIKITVPAHLEEFKISLGQSGRMALDIFSSTGRDIPELP 131  
 DB 81 NSTEDAVI---DLIKE-----ADFKIEL-----ETQTFD----- 106  
 OY 132 EREEGEGESELQNTAVSNMNPRLVELMRPEPSKLSIGIVGSGMGSRLSNGEVMRGIF 191  
 DB 107 -----KSDGQAGSDPRNGTGMQAKNKNQEGTNNNAS--GGGSGGGGCGGCMAGM- 158  
 OY 192 IKHVLDSPPAKNGTLKPGDRIVEAPSSQSESEPEKAPLCSVPPPPSAFAEMGSDHTQSS 251  
 DB 159 ---NRQSGMOKRNTFTFTASMKOKHSNVADEDEDETRDMTG-----RIRTEAGYEIDRAS 209

OY 252 A-----SKISODVDK--EDERGSWKNIRERGTLTGELHMELEKGS-GGLSLAGNK 303  
 DB 210 AGNCKLNQOEKDRKEQEDDEFQYTAKNKRYNMKK-DLRRIEYORDASKPLGLGLAGHK 268  
 OY 304 DRSRMSVPIVIGDIPGAAGKGRLOIADELLEINGQILYGRSHONASSIIRCAPSKYII 363  
 DB 269 DROKMACFVAGVDPRGALGSVD-IKPGDEIYEVNGNVLNKRNCHLNAVASVFNKVDG-KLV 326  
 OY 364 FIRKDAVNOAIVCGNAVEPLPSNSENLOKKEPEPTTSDAVALSSFRNVCHLELPK 423  
 DB 327 MITSRRKPNDEGMC---VKPI-----KKPPTASDETKFI-FDQFPARVQVVK 371  
 OY 424 DOGGGLGIAI---SEEDTLGVYIKSLTEHGAATDGRKVGQDILAVDDEIVGYPIEKF 480  
 DB 372 -EGFLGIVYIGKHAVESSGIFISDLREGSNAELAG-VKVGDMILAVQDVTLESNTDA 429  
 OY 481 ISLKTAKMYK---LTHAENPDSQAVPSAAGAAGEKKNSSOSLWVPQSGSPESIR 537  
 DB 430 TGLLRAGGVYTMILLTKE---EAIKAEKAEKKKEEKKPEEPQ----- 475  
 OY 538 NTSRSSTPAIRASDPATCPPIPGCETTIEISKRGGLSLVSGSD--TLGAFIIEHY 595  
 DB 476 -----EPATAEIKPNKKILILEKVEKPMGVIVCGKNNHVTTCVITH-VY 521  
 OY 596 EGAACKDGRMLWAGQDILEVNG--IDLKATYHDEAINVLRQTPOR-VRLTLYRDEAPYKE 652  
 DB 522 PEGQVAAKRLKIPHDIDINGTPIHVGSMITLAKVHOLFHTTEKAVTLVFRADPPELE 581  
 OY 653 EEVCDTLTIELOKKPKGKGLSLIVSGKRNDDGVFVSDIYKGGIADDPGRILIOGQDILLVNG 712  
 DB 582 K-----FNVDLMKRKAGKEGLSL--SPNEIGCTIADLIOGVPEDSKLQRODITTKENG 634  
 OY 713 EDVRNASEAVALLKCSLGTVTEVGRIK 742  
 DB 635 DALEGLPVOVCYALFKGANGKVMETVRPK 664

Search completed: July 12, 2001, 14:43:13  
 Job time: 226 sec

Fri Jul 13 15:00:14 2001

us-09-502-698-2.rpt

Page 12



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2001, 14:40:58 ; Search time 23.36 Seconds  
(without alignments)  
866.668 Million cell updates/sec

Title: US-09-502-698-2

Perfect score: 5085

Sequence: 1 MGNVSKESFERTINAKGN.....EEAVALKRKGVTLVWLS 1005

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 193259 seqs, 20144635 residues

Total number of hits satisfying chosen parameters: 193259

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

Issued\_Patents\_AA.\*  
1: /cgn2\_6/ptodata/2/1aa/5A.COMB.pep.\*  
2: /cgn2\_6/ptodata/2/1aa/5B.COMB.pep.\*  
3: /cgn2\_6/ptodata/2/1aa/6A.COMB.pep.\*  
4: /cgn2\_6/ptodata/2/1aa/6B.COMB.pep.\*  
5: /cgn2\_6/ptodata/2/1aa/PTCUS.COMB.pep.\*  
6: /cgn2\_6/ptodata/2/1aa/Backfile1.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description                           |
|------------|-------|-------------|--------|----|---------------------------------------|
| 1          | 509   | 10.0        | 2485   | 4  | US-09-290-640-46<br>Sequence 46, Appl |
| 2          | 500.5 | 9.8         | 2466   | 3  | US-09-080-855-12<br>Sequence 12, Appl |
| 3          | 500.5 | 9.8         | 2466   | 5  | PCT-US94-09943-2<br>Sequence 2, Appl  |
| 4          | 498   | 9.8         | 2465   | 2  | US-08-596-281-3<br>Sequence 3, Appl   |
| 5          | 498   | 9.8         | 2465   | 3  | US-09-100-804-3<br>Sequence 3, Appl   |
| 6          | 340   | 6.7         | 1112   | 3  | US-09-045-632-2<br>Sequence 2, Appl   |
| 7          | 340   | 6.7         | 1112   | 3  | US-09-045-632-3<br>Sequence 3, Appl   |
| 8          | 335   | 6.6         | 702    | 3  | US-09-045-632-15<br>Sequence 15, Appl |
| 9          | 335   | 6.6         | 1018   | 3  | US-09-045-632-16<br>Sequence 16, Appl |
| 10         | 335   | 6.6         | 1061   | 3  | US-09-045-632-32<br>Sequence 32, Appl |
| 11         | 334.5 | 6.6         | 610    | 1  | US-08-410-804-1<br>Sequence 1, Appl   |
| 12         | 334.5 | 6.6         | 610    | 1  | US-08-259-514-1<br>Sequence 1, Appl   |
| 13         | 334.5 | 6.6         | 610    | 2  | US-08-858-311-1<br>Sequence 1, Appl   |
| 14         | 334.5 | 6.6         | 1050   | 3  | US-09-045-632-50<br>Sequence 50, Appl |
| 15         | 331.5 | 6.5         | 1050   | 3  | US-09-045-632-49<br>Sequence 49, Appl |
| 16         | 329.5 | 6.5         | 604    | 3  | US-09-045-632-14<br>Sequence 14, Appl |
| 17         | 327   | 6.4         | 918    | 3  | US-09-045-632-21<br>Sequence 21, Appl |
| 18         | 327   | 6.4         | 961    | 3  | US-09-045-632-33<br>Sequence 33, Appl |
| 19         | 321   | 6.3         | 602    | 3  | US-09-045-632-20<br>Sequence 20, Appl |
| 20         | 288.5 | 5.7         | 861    | 3  | US-09-045-632-34<br>Sequence 34, Appl |
| 21         | 287   | 5.6         | 818    | 3  | US-09-045-632-25<br>Sequence 25, Appl |
| 22         | 273   | 5.4         | 507    | 3  | US-09-045-632-13<br>Sequence 13, Appl |
| 23         | 271   | 5.3         | 504    | 3  | US-09-045-632-19<br>Sequence 19, Appl |
| 24         | 266.5 | 5.2         | 642    | 3  | US-09-045-632-35<br>Sequence 35, Appl |
| 25         | 265   | 5.2         | 599    | 3  | US-09-045-632-28<br>Sequence 28, Appl |
| 26         | 258.5 | 5.1         | 502    | 3  | US-09-045-632-24<br>Sequence 24, Appl |
| 27         | 252   | 5.0         | 374    | 4  | US-09-091-405-2<br>Sequence 2, Appl   |

|    |       |     |      |   |  |
|----|-------|-----|------|---|--|
| 28 | 227   | 4.5 | 519  | 3 | US-08-997-445D-2<br>Sequence 2, Appl   |
| 29 | 225.5 | 4.4 | 283  | 3 | US-09-045-632-27<br>Sequence 27, Appl  |
| 30 | 219   | 4.3 | 498  | 3 | US-09-045-632-30<br>Sequence 30, Appl  |
| 31 | 219   | 4.3 | 541  | 3 | US-09-045-632-36<br>Sequence 36, Appl  |
| 32 | 217   | 4.3 | 284  | 3 | US-09-045-632-12<br>Sequence 12, Appl  |
| 33 | 214.5 | 4.2 | 407  | 3 | US-09-045-632-18<br>Sequence 18, Appl  |
| 34 | 209.5 | 4.1 | 1612 | 3 | US-08-545-860D-48<br>Sequence 48, Appl |
| 35 | 209.5 | 4.1 | 1612 | 5 | PCT-US94-04496-48<br>Sequence 48, Appl |
| 36 | 203   | 4.0 | 404  | 3 | US-09-045-632-23<br>Sequence 23, Appl  |
| 37 | 200   | 3.9 | 79   | 3 | US-09-100-804-27<br>Sequence 27, Appl  |
| 38 | 200   | 3.9 | 86   | 5 | US-08-545-860D-53<br>Sequence 53, Appl |
| 39 | 196.5 | 3.9 | 86   | 5 | PCT-US94-04496-53<br>Sequence 53, Appl |
| 40 | 196.5 | 3.9 | 198  | 4 | US-09-045-632-11<br>Sequence 11, Appl  |
| 41 | 190.5 | 3.7 | 1829 | 4 | US-09-157-420-1<br>Sequence 1, Appl    |
| 42 | 176.5 | 3.5 | 182  | 3 | US-09-045-632-29<br>Sequence 29, Appl  |
| 43 | 176.5 | 3.5 | 233  | 2 | US-09-151-611-1<br>Sequence 1, Appl    |
| 44 | 173   | 3.4 | 505  | 1 | US-08-123-161A-14<br>Sequence 14, Appl |
| 45 | 173   | 3.4 | 505  | 1 | US-08-483-278-14<br>Sequence 14, Appl  |

#### ALIGNMENTS

RESULT 1  
US-09-290-640-46  
Sequence 46, Application US/09290640  
Patent No. 6204055  
GENERAL INFORMATION:  
APPLICANT: Dean, Nicholas M.  
APPLICANT: Marcussen, Eric G.  
TITLE OF INVENTION: Antisense Compound Modulation of Fas Mediated Signaling  
FILE REFERENCE: ISPH-0351  
CURRENT APPLICATION NUMBER: US/09/290, 640  
CURRENT FILING DATE: 1999-04-12  
NUMBER OF SEQ ID NOS: 85  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 46  
LENGTH: 2485  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-290-640-46

|                       |                   |   |             |              |
|-----------------------|-------------------|---|-------------|--------------|
| Query Match           | 10.0%             | Score 509;  | DB 4;       | Length 2485; |
| Best Local Similarity | 22.8%             | Pred. No. 1.1e-34;  |             |              |
| Matches 243;          | Conservative 140; | Mismatches 371;   | Indels 314; | Gaps 37;     |
| OY                    | 162               | EPKSLIGISYVGRMGSRISNGEVMRGIFIKHYEDSPAGKNGTLKPGDRIVEAPSOE    | 221         |              |
| DB                    | 990               | EPPPTVALVELYKPSHOKSRSDAESLAGY---TKLNKSKVASLNRSEPRKRHESDSSI  | 1046        |              |
| OY                    | 222               | SEPERAPLCSVPPPPPSAFAEMGSDHTQSASAKISQDYDKDEDFGYSWKNIERYGTLLG | 281         |              |
| DB                    | 1047              | EDPGAYVLGMT-----MHSSGSSSQVPLKENDV-----LHKRMSIVSS            | 1086        |              |
| OY                    | 282               | ---ELHMELEK-GHSGIGLSIAGKKDSR--MSFYIGIDPNCAAGDKRGQIDELLE     | 335         |              |
| DB                    | 1087              | PERETTLVNLKRDARYGAGFQIIGGKMRDLGIFISSVAGGADLDGCLKPEDRLIS     | 1146        |              |
| OY                    | 336               | INGQILYRSHONASSIIRKCAPSKYIIFIRNKDQVNONMAVPCGAAVPELPSNSNLRK  | 395         |              |
| DB                    | 1147              | VANSVLEGVSHHALEIIONAPEDVTLVISQKEIKSVSPVPV---LTMENMYKK       | 1202        |              |
| OY                    | 366               | ETEPVTTSDAAVDISSRK-----VOHLELPK--DOGL--GIAISEEDTLGSIIRKS    | 445         |              |
| DB                    | 1203              | SS---YMDSAIDSSKDHHSRGLTNRHISNSPGGLREGSLSDSDRTESASLSQ        | 1258        |              |
| OY                    | 446               | LTEHVATDGLKKGDDILAVDDIIVGYPLEKISILKTRAKMYKLTJIAENDSQAV      | 505         |              |
| DB                    | 1299              | SOVNVFFAS---HIGDQTW---QESOHGSPSPVIS-----KATEKETFDSDNSKTRK   | 1306        |              |
| OY                    | 506               | PSAAGAASGEK-----NSQSILVPOSGSPESIRTRRSSTPAIFASDPATC          | 555         |              |

Db 1307 PGISDVTYDSDRGSDMDDEATYSSQDHQTPKQ---ESSSSVNTSNKMNFTSSSPK- 1362  
 QY 556 PIPGCEITTEISKRGIGLSIVGSDPTLL--GAFITHEVEGACCKDRGLMAGDOIL 613  
 Db 1363 ---PGDIFEVFLANKNDNSLGSIVGVNTSVRHGIIYKAVIPQGAESDRIRKGRVL 1419  
 QY 614 EVNGIDLKATKTHDAINVLROTQORVRLTYRDEAPYKEEV-----C----- 656  
 Db 1420 AVNGVSLGGAHAKOAVETLRNTGQVNHLLLEKQSPTSKEHVPTPQCTLSDNAQGGP 1479  
 QY 657 -----DLTILELQKKPKGGLS-----IVGRKNDGVVSDIVK 691  
 Db 1480 EKVAKTQVKDYSVTEENTEVLFKN--SSGLGPFSSREDNLLPEQINASTIVYKKLEP 1538  
 QY 692 GGIADPDRLGDOILLVNGEDVRNASSDEAVALLKCSLGTVLEVRIGAG----- 744  
 Db 1539 GQPAAESKIDVGVILKVNASLKLGSQDEVIALKGTAEVLLLCRPPGVLPEDT 1598  
 QY 745 ---PFHSE-----RRPSQTSQVSEGS-- 762  
 Db 1599 ALNPLPSQAVLPNSSKSSQPCVEQSTSSDENEMSKKQCKSPSRDYSDDSGS 1658  
 QY 763 ---LSSFTF-----PLSGSST----- 775  
 Db 1659 GEDDLVTAPANISNSTWSSALHQTLSNMVSOAQSHNEAPKQEDTICTMFTYPOKIPNKP 1718  
 QY 776 -----SESLESSK-----KNALASEI 792  
 Db 1719 EFEDSNPSLPDPMAFGOSYQPOSESASSMDKXHHIHSEPTROENMTPLKNDLENHL 1778  
 QY 793 QGLR-----IVEMKKGPDTLSIGSIAG--VGSPLGDVPFIAMHPTGVAAOTOKLRY 844  
 Db 1779 EDFEVELLITLTKSEKSLGFTVTKGNQRIQCYVHDV-----IQDP--AKSDGRKLP 1830  
 QY 845 GDRIVTICGTBEQMTHTQAVNLLKNASGSIEMQVAVGADVVTGHHQF-PASSLSFT 903  
 Db 1831 GDRILKVNDDVTMTHTDAVNLRAASKTVRL-----YIGVLELPRJPM- 1877  
 QY 904 GLTSTIFODDLGPPCKSITLERGPDGLGFSIVGVGSPHGLDPIYKTVFAKAASED 963  
 Db 1878 -----PHLPDITLTCNKELEGSLOGGHSLL--QVVIYISDINPSVAIE 1922  
 QY 964 GRLRGDOIIAVNGOSLEGVTHEA-----VAIKRTKGTVTLM 1002  
 Db 1923 GNLDLIVIHVNGVSTQGMTLEEVNRALDMKSLPVLVKATRNDLPVY 1970  
 RESULT 2  
 US-09-080-855-12  
 : Sequence 12, Application US/09080855A  
 : Patent No. 6083721  
 : GENERAL INFORMATION:  
 : APPLICANT: Saras, Jan  
 : APPLICANT: Franz, Petra  
 : APPLICANT: Aspenstrm, Pontus  
 : APPLICANT: Hellman, Ulf  
 : APPLICANT: Genez, Leonel Jorge  
 : APPLICANT: Heidlin, Carl-Henrik  
 : TITLE OF INVENTION: PARC. A GTPASE ACTIVATING PROTEIN WHICH INTERACTS WITH PTPPLI  
 : FILE REFERENCE: 10461/7030  
 : CURRENT APPLICATION NUMBER: US/09/080,855A  
 : EARLIER FILING DATE: 1998-05-18  
 : EARLIER APPLICATION NUMBER: 08/805,583  
 : NUMBER OF SEQ. ID NOS: 39  
 : NUMBER OF SEQ. ID NOS: 39  
 : SOFTWARE: FastSeq for Windows Version 3.0  
 : SEQ ID NO 12  
 : LENGTH: 2466  
 : TYPE: PRT  
 : ORGANISM: Homo sapiens  
 : US-09-080-855-12

Query Match 9.8%; Score 500.5; Db 3; Length 2466;  
 Best Local Similarity 22.2%; Pred. No. 6.2e-34;  
 Matches 255; Conservative 146; Mismatches 396; Indels 351; Gaps 41;  
 QY 142 ELQATASNNQPRVVELMREPSLSLSIYGGNGMSRLSNG----- 184  
 Db 868 QLOMRARQSNODADIERASFRSLNLAESVGRFMGRALSTGLASTLNKLAIVRLSV 927  
 QY 185 --EVMRGIFIKHV-----LEDSPAGKNGTL---KPGD--RIYAPQSESEPEKAPLCS 231  
 Db 928 QAEILKRLSCSELLYQPLONSSKEKNDKASWEKPREMSSTYNDLSQASLYPIRKNNIV 987  
 QY 232 VPPPPSAFAEM-GSDHTQSSAS-----KISQVDKED 263  
 Db 988 NMEDPPQVAVELVYKPSHOMSRSDAESLACVTKLNKSVASLNRSPERRKHESSSIE 1047  
 QY 264 EFG--YSKNIREKRYGLTG---ELNHIELEK-GHSLGLSLAGNKRDR--MSVFIYGI 315  
 Db 1048 DPGQAVYLDVLRKMSYVSSPEREITLVNLKDKAKYGIGFOIIGEEKMGRDLGIFISV 1107  
 QY 316 DPNCAAGKDRLOIADLELLEINGOILYGRSHONASSTIKCAPSKVKTIFIRNKDAVNOMA 375  
 Db 1108 APGGPADHGLKPGDRLISVNSVSLGVSHHAAIEILQNAPEDTVLTISQPKKISKVP 1167  
 QY 376 VCPGNAAVEPLPSNSENLOKNETEPTVTSDAAYDLASSFKN-----YQHLLEPK--DOGG 427  
 Db 1168 STPYH-----LTNEMKNMTMKSS-----YMDSALIDSSKDHIMSGTLRHISENFGSPSG 1219  
 QY 428 L--GIAISEEDTSLGVYIKSLTEHGVAAIDRLKVGDOILAVIDEIVGVPIKEFISLKL 485  
 Db 1220 LRESLSDSDRTESASISQSVQVNGFPAS---HAGDQTV---QESOHGSPSPVIS--- 1269  
 QY 486 TAKTIVTLTHAENPDSQAVPSAAGAASGEKK-----NSQSILAMPQSSPEPES 535  
 Db 1270 --KATEKETFDLSQOSTKPKPGISDVTYDSDRGSDMDDEATYSSQDHQTPKQ---ESSS 1324  
 QY 536 IRTSRSTPAIFASDPAATPPIPCETTEIISKRTGLSIVGSDPTLL--GAFITHE 593  
 Db 1325 SVNTSNKMNFTFSSSPK---PGDIFEVFLANKNDNSLGSIVGVNTSVRHGIIYKVA 1380  
 QY 594 VYEGGACCKGRRLMAGDOILLEVNGIDLRKATHDAINVLQTPQORVRLTYRDEAPYKEE 653  
 Db 1381 VTPQGAESDGRIRKGRVILAVNGVSLGGAHAKOAVETLRNTGQVNHLLLEKQSPSK 1440  
 QY 654 EV-----C-----DLTILELQKKPKGGLS----- 674  
 Db 1441 HVPYTPQCTLSDNAQGGPPEKAVKTTQVNDYSVTEENTEFVLFKN--SSGLGPFSSRE 1499  
 QY 675 ---IVGRKNDGVVSDIVYKGIADPDRLGDOILLVNGEDVRNASSDEAVALLKCSL 731  
 Db 1500 DNLIPEQINASTIVYKRLFLGQPAAESKIDVGVILKVNASLKLGSQDEVIALRGTA 1559  
 QY 732 GTVTVLEVRIGAG-----PFHSE----- 749  
 Db 1560 PEVFLLLCRPPGVLPEDITALLTPLOSQAVLPNSSKSSQPCVBOSTSSDENEMSK 1619  
 QY 750 ---RRPSQTSQVSEGS-----LSSTF-----PL 770  
 Db 1620 SKKCKCKSPSRDYSDDSGGEDDLVTAPANISNSTWSSALHQTLSNMVSOAQSHNEAPK 1679  
 QY 771 SGSST-----SESLESSK----- 784  
 Db 1680 SOEDTICTMFTYPOKIPNKEFEEDSNPSLPDPMAFGOSYQPOSESASSMDKXHHIH 1739  
 QY 785 -----KNALASEIQGLR-----IVEMKKGPDTLSIGSIAG--VGSPLGDVP 824  
 Db 1740 SEPTROENMTPLKNDLENHLEDELEVELLITLTKSEKASLGFTVTKGNQRIQCYVHDV- 1798  
 QY 825 IFIAMHPTGVAAOTOKLRYGDRIVTICGTBEQMTHTQAVNLLKNASGSIEMQVAVGAD 884  
 Db 1799 ---IQDP---AKSDGRKLPKGRILKVNDDVTMTHTDAVNLRAASKTVRL----- 1844  
 QY 885 VSVVTGHHQF-PASSLSFTGLTSTIFODDLGPPCKSITLERGPDGLGFSIVGVGSP 943

Db 1845 ---VIGRVLPRIPML-----PHLLPDTITLCKNEELGFSLCGGHDL 1885

Qy 944 HGDLPYKTVFAKGAASEDRKLRGDDQIIAANGSLSEVTHEEA-----VAILKR 994

Db 1886 Y--QVYVYIDINPRVAIAEGNLOLDLVHYVNGVSTOGMTLEEVNRALDMSLPISYLKA 1943

Qy 995 TKGTVTLM 1002

Db 1944 TRNDLPV 1951

RESULT 3  
PCT-US94-09943-2

Sequence 2, Application PC/TUS9409943

GENERAL INFORMATION:

APPLICANT:

APPLICANT:

APPLICANT:

APPLICANT:

APPLICANT:

TITLE OF INVENTION: PRIMARY STRUCTURE AND FUNCTIONAL

TITLE OF INVENTION: EXPRESSION OF NUCLEOTIDE SEQUENCES FOR NOVEL PROTEIN

TITLE OF INVENTION: TYROSINE PHOSPHATASES

NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:

STREET: 600 ATLANTIC AVENUE

CITY: BOSTON

STATE: MASSACHUSETTS

COUNTRY: USA

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: IBM PC compatible

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US94/09943

FILING DATE: 01-SEP-1994

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/115,573

FILING DATE: 01-SEP-1993

ATTORNEY/AGENT INFORMATION:

NAME: TWOMEY, MICHAEL J.

REGISTRATION NUMBER: P-38,349

REFERENCE/DOCKET NUMBER: L0461/700000

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617/720-3500

TELEFAX: 617/720-2441

TELEX: 92-1742 EZEKIEL

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 2466 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

PCT-US94-09943-2

Query Match 9.8%; Score 500.5; DB 5; Length 2466;

Best Local Similarity 22.2%; Pred. No. 6.2e-34;

Matches 255; Conservative 146; Mismatches 396; Indels 351; Gaps 41;

Qy 142 ELQNTAAYSMNPNRVELMREPSKLSIGISVGRGKMSRLSNG----- 184

Db 868 QLOMAROSNODADIERASFSRLNLOAESVGFNMGRASITGLASSTLKLAVPLSV 927

Qy 185 ---EWKGRIFIRKIV-----LEDSPAGKNGTL-----KPGD--RIEAPSOSESSEPEKAPLDS 231

Db 928 QAEILKRLKSCSELSTYQPLQNSKSKENKAKASWEKPREMSKYHDLQASLYPHRRKNVIV 987

Qy 232 VPPPPSAFAEM-GSDHTOSSAS-----KISQDVDKED 263

Db 988 NMEPPQVYAEVLGKPSHQMSRDAESLAGVTKLNNKSVASLNSPERKKHEDSSSIE 1047

Qy 264 EFG--YSKKNIERYGTLTG---ELHMIELK-GHSGGLSLAGKNDNR--MSVFIYGI 315

Db 1048 DGOAYVLDVLRKMSIYSSPEREITLVNLKKDARYGGLGFOIIGEGKMRDLGIFISSV 1107

Qy 316 DNGAAGKGRLOADELLEINGQILYGRSHQNASIIKCAFSKYKILIRKDDVNOA 375

Db 1108 AFGPRDFHGLKPGDRLIVSNVSLGESHHAALIEIIONAPEVTLVYQREKITSVP 1167

Qy 376 VCPGNAVEPLPNSSENLOKETEPTVTTSDAAVDSLSEFN-----VOHLLEPK--DOG 427

Db 1168 STPVH---LTNEMKNYMKSS---YQDSALIDSSSKOHMSRGLRHSNSRGPSCG 1219

Qy 428 L--GIAISEEDTLGVIYIKSLTEHGYAATDGLKYGDDQILAVALDEIVYGPLEKTIISLK 485

Db 1220 LREGSLSSQDSRTEASLSQSQVNGFFAS---HLGDQW---QESQHSPPSVYS--- 1269

Qy 486 TAKMTVYKLTIAENPDSQVPSAAGAAGEKK-----NSSOSLWVPOGSPPEPS 535

Db 1270 --KATEKETFTDSNOSKTKKPGISDVTYSDRGSDMDENATYSSQDHTPKQ---ESS 1324

Qy 536 INTRSSSTPAIFADDPATCIIPCCETIIEISKRTGLSIGSVGSDTL--GAFIIE 593

Db 1325 SVNTSNKNMFKTFSSPPK---PGDIFEVLLAKNDNSIGISVGTGVTNVHRGIIYKA 1380

Qy 594 VYEEGAACKDGLMAGDQILEVNGIDLKRAIIDEALNVLRQTPQVRLTLVDEAPYKE 653

Db 1381 VIPOGAESDGRIRHKGRVLAANGVSLGEGATHKAVETLRNGOVVHLLLEGQSTSK 1440

Qy 654 EV-----C-----DITLIELOKKRKGIGLS--- 674

Db 1441 HVPVTPQCTLSQNMAGQPEKVKTKTYQKDSFYTEENTFEVKLFKN--SSLGFSFSRE 1499

Qy 675 ---IVGRNDGTGVYSDIVYKGIADPGRILGDDQILLVNGEDVRNASEAVAAALKCSL 731

Db 1500 DNILPEQIINASIVRYKILPAGOPAAESGKIDVDVILKNGASLKGSLQOEYISALRGTA 1559

Qy 732 GTVTEVGRIKAG-----PHSE----- 749

Db 1560 PEVFLLCRPPGVLPETIDTALLPLQSPAQVLPNSSKDSQSCVEQSTSDENEMSDK 1619

Qy 750 ---RRPSOTQOVSEGS-----LSPTF-----PL 770

Db 1620 SKKCKSPSRBRSYSDSSSGGDDLVTA PANISNWSALMOTLSNMYSQAQSHHEAK 1679

Qy 771 SGSSST-----SESLESSK----- 784

Db 1680 SOEDTICTMFFYQKIPNKPFEEDSNPSPLPPDMAFGQSYQPOSESASSSMDKYHIHI 1739

Qy 785 ---KNMLASRIQGR---TYEMKKGPTISIGISAG---VGSPLGDVP 824

Db 1740 SEPTROENNTPLKNDLENLEPELEVEELLITLILSEKASISLPTVYKGNORIGCYVHDV- 1798

Qy 825 IFIAMHPTGVAAGTQKLVAGDRIYITCGSTEGMTHTQAVNLKKNASIEIQVAGGD 884

Db 1799 ---IDP---AKSDGRLEPGDRLIKVNDTDTNMHTAVNLDRAASTIVNL----- 1844

Qy 885 VSVVTGHOE-PASSLSLTGTLSTISIFODDLGPPCKSITLERGPDGLGFSIVGYSGP 943

Db 1845 ---VIGRVLPRIPML-----PHLLPDTITLCKNEELGFSLCGGHDL 1885

Qy 944 HGDLPYKTVFAKGAASEDRKLRGDDQIIAANGSLSEVTHEEA-----VAILKR 994

Db 1886 Y--QVYVYIDINPRVAIAEGNLOLDLVHYVNGVSTOGMTLEEVNRALDMSLPISYLKA 1943

Qy 995 TKGTVTLM 1002

Db 1944 TRNDLPV 1951

RESULT 4  
US-08-596-291-3

Sequence 3, Application US/08596291  
Patent No. 5821075  
GENERAL INFORMATION:  
APPLICANT: GONZ, LEONEL JORGE  
APPLICANT: SARAS, JAN  
APPLICANT: CLAESSON-WELSH, LENA  
APPLICANT: HELDIN, CARL-HENRIK  
TITLE OF INVENTION: PRIMARY STRUCTURE AND FUNCTIONAL  
TITLE OF INVENTION: EXPRESSION OF NUCLEOTIDE SEQUENCES FOR NOVEL PROTEIN  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: WOLF, GREENFIELD & SACKS, P. C.  
STREET: 600 ATLANTIC AVENUE  
CITY: BOSTON  
STATE: MASSACHUSETTS  
COUNTRY: USA  
ZIP: 02210  
COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/596,291  
FILING DATE: 09-AUG-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/115,573  
FILING DATE: 01-SEP-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: GATES, EDWARD R.  
REGISTRATION NUMBER: 31,616  
REFERENCE/DOCKET NUMBER: L0461/7000  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617/720-3500  
TELEFAX: 617/720-2441  
TELEX: 92-1742 EZEKIEL  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2465 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-596-291-3

Query Match 9.88; Score 498; DB 2; Length 2465;  
Best Local Similarity 21.98; Pred. No. 1e-33;  
Matches 252; Conservative 150; Mismatches 389; Indels 362; Gaps 41;

QY 142 ELQNTAVSNMNPRLVLRPEPSKLSIGISVGRGMSRLSNG-----184  
DB 868 QLRHARSNDADIDIEASFRSLNLAQESVRGIMGRAISTGSLASLTUNKLAVRLSV 927  
QY 185 --EVMRGIFIRHV-----LEDSPAGKNGTL---KPGD--RIVEAPSOSESEPEKAPLCS 231  
DB 928 QAEILKRLSCSELQLYOPLQNSKKNKDKASWEKPREMSKSYHDLQSASLYPHRKNIY 987  
QY 232 VPPPPSAFAM-GSDHOSAS-----KISDDVDEKED 263  
DB 988 NMEPPPOVALVGRKPSHQMSRSDAESLAGYTKLNNSKSVASLNKSPERRKHESDSSIE 1047  
QY 264 EFG--YSKNIRERYGLTG---ELHMLELK-GHSGIGLS-LAGNKDRSHMSYFIYID 316  
DB 1048 DPGQAVYVDLHKRMSVSSPEREITLVNLKKDKAYGIGRPIIIGEKMETDGLGFISSVA 1107  
QY 317 PNGAAGKDRLOLADLELNGQLLYGRSHQMSAIIKCAPSKYIIFIRNKDAVNOMAV 376  
DB 1108 PGGADEFGCKLPDDRILSVNSVLEGVSHAATIELIQNADEDTLVISQPKERIKSVPS 1167  
QY 377 CPGNAVEPLPNSNENLQNKETEPYVTSDAVNDLSSFRN-----VQHLLEPK--DQGL 428

DB 1168 TPVH-----LTNEMKNYMKSS-----YMODSAIDSSKDHMSRGLRLHISENSFGPSGGL 1219  
QY 429 --GIAISEEDLSGVIIKSLTEHGVAANTDGRKXGDDOILANDDIIVGYPIERKISLTKT 486  
DB 1220 REGSLSDOSTRESASISQSOVNGFFAS---HLGDTW---QSOHGSFSPSYIS----1288  
QY 487 AKMTVKLTIIHAENPDSOAVPSAAGAAGEKK-----NSSQSLAVPOSGSPERESI 536  
DB 1269 -KATEKEFTDSNOKSKTKKPEISDVTDSGDSDMDENATYSSQDHPOTPO---ESSSS 1324  
QY 537 RNTSRSSSTPAIFASDPATCPPIIPCEETIETISKRTGLGSLVSGSDTL--GAFITHEV 594  
DB 1325 VNTSNKNMFKTFSSSPK---PGDIEVELAKNDNSLGISVTVGVTSVYHGGIYKDV 1380  
QY 595 YEEGACKDRMLAGDDOILEVNGIDILRKATHEAIVNLRPQVVRVLTIVDEAPYKEE 654  
DB 1381 IPQGAESDGRHKHGRVLAIVGSLBGAITHKQAVETLRNTGVVHLLLEKGGSPSKEH 1440  
QY 655 V-----C-----DTLIELOKRPKGLS-----674  
DB 1441 VPVTPQCTLSQNAQGOGEPEKVKTKTQKDYSPVTEENTFEVKLFKN--SSGLGFSFRED 1499  
QY 675 --YGRKNDVGFVSDIVKGIADPDGRLOGDQIILVNGEDVYNAQEAVALKCSLG 732  
DB 1500 NLIPEQINASIVRVKTLFAGOPAESGKIDVDVILKVNGLSLGSLQOEYISALRGTA 1559  
QY 733 TVTLEVGRIKAG-----PFHSE-----749  
DB 1560 EVFLLICRPPGVLEPIDTALTPLQSPAQVLPNSSKDSQSPCEVOSTSDENEMSDKS 1619  
QY 750 ---RRPSQTSQVSEGS-----LSFTF-----PLS 771  
DB 1620 KKQCKSPSRSDSYSDSGCEDLVTPANISNSTWSALHOTLSNNVSOAOSHHEAPKS 1679  
QY 772 GSST-----SELSVSK-----784  
DB 1680 QEDTICTMFTYPOKIPNKPREDNSPBLPDPMAPOGSIOPQSSASSSMDKHHIHS 1739  
QY 785 -----KNALASEIOGLR-----YEMKGPYDLSISIAGC---VSPPLGDPVI 825  
DB 1740 EPTROENMTPLAKNLEHLEDFELEVELLITLISEKASLGFTYTKQNRIGCVHVV--1797  
QY 826 FIAMHPGVAAQTKRLRGVDRIVTICGSTEGTTHPOAVNLKNASGSIEM-----QV 879  
DB 1798 ---IODP---AKSDGRILKPGRLIKVNDTQVNTHTDAVNLRAASKTVLVIYGRPRI 1851  
QY 880 VAGGDV--SVYTGHHQEPASSLSFTGLSTSIIPQDDLGPQCKSITLERGPDGCFISYG 938  
DB 1852 TQNTNVASFATGH-----KLTCKNEELGFSLCG 1879  
QY 939 GYGSPPHGLPIYVTVFAKGAASEDRGLRKQDQIIVANGOSLEGVTHEEA-----V 989  
DB 1880 GHSDLY--QVYISDIDNPRVAALIEGNQLDDVHYVNGVSTQGMTLEEVNRLADNLSLP 1937  
QY 990 AILKRTKGYTIM 1002  
DB 1938 LVLKATRDLPVV 1950

RESULT 5  
US-09-100-804-3  
Sequence 3, Application US/09100804  
Patent No. 6066472  
GENERAL INFORMATION:  
APPLICANT: GONZ, LEONEL JORGE  
APPLICANT: SARAS, JAN  
APPLICANT: CLAESSON-WELSH, LENA  
APPLICANT: HELDIN, CARL-HENRIK  
TITLE OF INVENTION: PRIMARY STRUCTURE AND FUNCTIONAL  
TITLE OF INVENTION: EXPRESSION OF NUCLEOTIDE SEQUENCES FOR NOVEL PROTEIN  
NUMBER OF SEQUENCES: 34  
CORRESPONDENCE ADDRESS:

ADDRESSEE: WOLF, GREENFIELD & SACKS, P.C.  
 STREET: 600 ATLANTIC AVENUE  
 CITY: BOSTON  
 STATE: MASSACHUSETTS  
 COUNTRY: USA  
 ZIP: 02210  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 OPERATING SYSTEM: IBM PC compatible  
 SOFTWARE: Patent Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/100, 804  
 FILING DATE:  
 CLASSIFICATION:  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/596,291  
 FILING DATE: 09-AUG-1996  
 APPLICATION NUMBER: US 08/115,573  
 FILING DATE: 01-SEP-1993  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: PCT/US94/09943  
 FILING DATE: 01-SEP-1994  
 ATTORNEY/AGENT INFORMATION:  
 NAME: GATES, EDWARD R.  
 REGISTRATION NUMBER: 31,616  
 REFERENCE/DOCKET NUMBER: L0461/7003  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 617-720-3500  
 TELEFAX: 617-720-2441  
 TELEX:  
 INFORMATION FOR SEQ ID NO: 3:  
 SEQUENCE CHARACTERISTICS: 3  
 LENGTH: 2465 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-09-100-804-3

Query Match 9.8%; Score 498; DB 3; Length 2465;  
 Best Local Similarity 21.9%; Pred. No. 1e-33;  
 Matches 252; Conservative 150; Mismatches 389; Indels 362; Gaps 41;

Db 1269 KATEKEPTDSDNOSKTKKRGISDVTDYSDRSDSDMDENATYSSSDHQHPKQ---ESSSS 1324  
 QY 537 RNTSSSTPAIRASDPANCPPIPGCEETIEISKRTGLGLSTVGSSTLL--GAPITHEV 594  
 Db 1325 VNTSNKMNFKETSSSPK---PGDIFEEELAKNDNSLIGSTVGCVNTSVRRHGIVYKV 1380  
 QY 595 YEEGAACKDGRIMAGDQILEVNGIDLRKATHEAIVNROTQORRLTYRDEAPYKKEE 654  
 Db 1381 IPGGAESDGRILHKDRLVAVNGVSLGATKQAVETLNTQVYVHLLLEKQSPYSKEH 1440  
 QY 655 V-----C-----DTLTLEQKKPKGGLS----- 674  
 Db 1441 VPTPQCCTLSDQNAQGQGEKVKTTQVKDYSPVTEENFEVKLEKKN--SSGLGFSRSD 1499  
 QY 675 --IVGRNDTVGVSDIYKGIADPDGRLIGDQILLVNGEDVRNNAQPAVALLKCSLG 732  
 Db 1500 NLIPQINASIYRVKFLFAGQPAESGKIDVGVILKVGASLKGISQOEVIASALRGTA 1559  
 QY 733 TVTLEVGRIKAG-----PFHSE----- 749  
 Db 1560 EVFLLCRPPGVLPIDITALLTPIQSPAQVLPNSSKSSDPSVCYQSTSSDENESDKS 1619  
 QY 750 ---RRPQTSQVSEGS-----LSFTF-----PLS 771  
 Db 1620 KKQCKSPSRSDYSDDSSGSGEDDLYTAPANISNTWSSALHOTLSNMVSOAQSHHAPKS 1679  
 QY 772 GSST-----SESLSSSK----- 784  
 Db 1680 QEDTICTWFEYPOKIPNKFEDSDNPSPLPDPMAQSOYQOPSSSSSMKHYHHTS 1739  
 QY 785 -----KNALASETOGLR-----IVEMKKGTSDLSISING-----VSPGLDVP 825  
 Db 1740 EPTRODMTPLKNDLENHLEDELEVELLITLIKSEKASLGFTYKGNORICYVADV-- 1797  
 QY 826 FLAMHPGTVAQOTOKLRYGDRIVITIGTSTEGMTHTQAVNLKNAQSIEM-----QV 879  
 Db 1798 ---IQDP---AKSDGLKGRDLIKVNDVYVNMHTDANVNLRAASKYRVLVIGVPR 1851  
 QY 880 VAGGV-SVVTGHQHOEPASSLSFTGLSTSLFODDLPPOCKSTLERGPDGLGFSIVG 938  
 Db 1852 TONTNVAEPATCH-----KLTCKNKEELGFSLCG 1879  
 QY 939 GIGSPHGDPIYKTYFAKGAASEGRLKRGDIIIVANGOSLEGYTHEA-----V 989  
 Db 1880 GHDSTL--QVYVTSIDNIPRSVAIREGNIQLDLYIHVNGVSTQGMTELEVNRALDMSLPS 1937  
 QY 990 AIHKRTKGTVTL 1002  
 Db 1938 LVLKATRNDLPV 1950

RESULT 6  
 US-09-045-632-2  
 Sequence 2, Application us/09045632  
 Patent No. 6001575  
 GENERAL INFORMATION:  
 APPLICANT: HUGANIR, Richard L.  
 APPLICANT: DONG, Hualing  
 TITLE OF INVENTION: THERAPEUTIC USES OF GRIP AND  
 NUMBER OF SEQUENCES: 105  
 CORRESPONDENCE ADDRESS:  
 ADDRESSSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
 STREET: 130 Water Street  
 CITY: Boston  
 STATE: MA  
 COUNTRY: USA  
 ZIP: 02109  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 OPERATING SYSTEM: PC-DOS/MS-DOS

```

SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/045,632
FILING DATE: 19-MAR-1998
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/041,016
FILING DATE: 19-MAR-1997
ATTORNEY/AGENT INFORMATION:
NAME: Corless, Peter F.
REGISTRATION NUMBER: 33,860
REFERENCE/DOCKET NUMBER: 48147/1699-CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1112 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-045-632-2

```

```

Query Match 6.7%; Score 340; DB 3; Length 1112;
Best Local Similarity 22.6%; Pred. No. 1.5e-20;
Matches 173; Conservative 90; Mismatches 265; Indels 238; Gaps 22;

```

```

QY 413 EKNVHLEPRDQ-GIGIAISEDTLSGV-----IKSLTEHVAATGRKLVGDI 464
DB 48 FKSGTVELMKEGTTLC-----TVSGIDKDGKPRVSNLRGGIAARSDQLDVGYI 101
QY 465 LAVDEIVGVPIKFIKSLTAKMTVLTTHAENPDQAVPSAAGASGKKNSQSILM 524
DB 102 KAVGINLAKRHEIISLKNVGERVYLEVEYELP----- 137
QY 525 VPQSGSPESIRNNTSSSTPAIRASDPATCPPIPGCTTIEISKGRG--LGSIYGS 582
DB 138 -----PVSIGSS-----VME-----RTVEVTLHKEGMTFGVIRGGA 170
QY 583 D---TLGAFIHEVEGACAKDGRMAQDQILEVNGIDLRKATHDAINVLRQTPQV 639
DB 171 HDNRKSPVITCYRPGGPRDRECTIKPRDLISVDGIRLGLTHAFAISILKQCGOA 230
QY 640 RLTL-YRDEAPYKEEVCDTLTIELQKPKGKLSIYVK--RNDTGVFVSDIYKGIAD 696
DB 231 TLLEIVDSANDSVATASGPLVEAKTPGASLGLVLTSCVKQYVIDIKISASIA 290
QY 697 PDGRLIQGDQILLVNGEDVRNASEAVALLKCSLGVTVLEV-----GRK-AGPFH-- 747
DB 291 RCGALHVGDIHLSIDGSMETLAETQFLGNTTDQVKEILPHHOTRLALKGPDIKI 350
QY 748 -----SERPSQ 754
DB 351 QRSRDLPMWPMWASSQCVHTNNHHNPHHDPCHVPALGFPKALTPNSPPMVAWSSSS 410
QY 755 TSOVSEGLSSTFPLSGSSTS-----ESLESSKKNALASEIQLGLTV 798
DB 411 MSAYSLSLNNGTLPRSLYSTSPRGTMWRRLKKKDFKSSLSLSTYGLAGQVHHETT 470
QY 799 E--KKGPTDLSGISAGV--GSPGLGVPIITAMHHPGVAAGQOKLRVGRIVYTIGTS 855
DB 471 EVVLADPVGFGIOLGQSVFATETLSPPPLISYIEADSPERCGVLQIDRVAIAINIP 530
QY 856 TEGHTHQAVALLNKASS-----IEMQVYAGDVSVYVGHNEPASPSSLSFGLNITS 909
DB 531 TEGSTFEBANOLKRDSSITTSVLTLEIEPDAESVTPSSGTFHVKLPKHVEL-GITIS 589
QY 910 IFODDLGP-----GCK--- 921
DB 590 PSSRKPDPPLVSIIDIKKGSVAHRTGTLELGDKLLAIDNIRLSDCSMEDAVQILOGCEDLV 649

```

```

QY 922 -----SITLER--GPDGLGSIYGVGSPHDDLPYVKTFAK 957
DB 650 KLRKRDENSDDEQSSCAITYVELKRYGGPLGITIS-----GTPEPFDPIIISLTKG 704
QY 958 GAASEDKLRKGDQIIANVNGSLGCVTHEEVAVALIKRKGVTVLMV 1003
DB 705 GLAERTGAIHIGDIRLAINSSSLKGPLSEDIHLLQMAGETVTVTKI 750

```

# RESULT 7

```

US-09-045-632-3
Sequence 3, Application US/09045632
Patent No. 6001575
GENERAL INFORMATION:
APPLICANT: Huganir, Richard L.
APPLICANT: Dong, Hualing
TITLE OF INVENTION: THERAPEUTIC USES OF GRIP AND
NUMBER OF INVENTION: GRIP-RELATED MOLECULES
NUMBER OF SEQUENCES: 105
CORRESPONDENCE ADDRESS:
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/045,632
FILING DATE: 19-MAR-1998
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/041,016
FILING DATE: 19-MAR-1997
ATTORNEY/AGENT INFORMATION:
NAME: Corless, Peter F.
REGISTRATION NUMBER: 33,860
REFERENCE/DOCKET NUMBER: 48147/1699-CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1112 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-045-632-3

```

```

Query Match 6.7%; Score 340; DB 3; Length 1112;
Best Local Similarity 22.6%; Pred. No. 1.5e-20;
Matches 173; Conservative 90; Mismatches 265; Indels 238; Gaps 22;

```

```

QY 413 EKNVHLEPRDQ-GIGIAISEDTLSGV-----IKSLTEHVAATGRKLVGDI 464
DB 48 FKSGTVELMKEGTTLC-----TVSGIDKDGKPRVSNLRGGIAARSDQLDVGYI 101
QY 465 LAVDEIVGVPIKFIKSLTAKMTVLTTHAENPDQAVPSAAGASGKKNSQSILM 524
DB 102 KAVGINLAKRHEIISLKNVGERVYLEVEYELP----- 137
QY 525 VPQSGSPESIRNNTSSSTPAIRASDPATCPPIPGCTTIEISKGRG--LGSIYGS 582
DB 138 -----PVSIGSS-----VME-----RTVEVTLHKEGMTFGVIRGGA 170
QY 583 D---TLGAFIHEVEGACAKDGRMAQDQILEVNGIDLRKATHDAINVLRQTPQV 639

```

```

Db 171 HDRNRKSRPVVITCVPRGPGDDREGTIKRGDRLLSVDIGRLITTHAEMSLKCGQGA 230
Qy 640 RLTL-YRDEAPRKEEVEDTLTIELQKRGKGLSIVK--RNDTGVVSDIVGGIAD 696
Db 221 TLLIYDVASMDSVATASGPLVEAKTPGASLVALTTSVCCNKQVITIDIKASAD 290
Qy 697 PDGRLIOGDIILVNGEDVRNASQEAVALLKCSLGTVLEY-----GRK-AGPFH--- 747
Db 291 RGAHLVGDHILSIDGTSMETCTLAETQFLGNTDQVKLELPHHQRLALKGDDHVKI 350
Qy 748 -----SERPSQ 754
Db 351 QRSRQLPMDPWASQCSVHTNNHHNPHHCHRVALGFPKALTPNSPRAWSSSPTS 410
Qy 755 TSOVEGSLSTFPPLSGSSTS-----ESLESSKKNAASEIOLKRV 798
Db 411 MSAYSLSLNMGTLPRLSTSPRCTMMRRLLKKDFKSSLSLASTVGLAQVHTET 470
Qy 799 E--MKKGPDSLSIGAGV--GSPLDGVPIFTAMHPTGVAOQOKLVRGDRIVTICGS 855
Db 471 EVVLADPYTGIGIOLQSGVFATETLSFPLISYLEADSPARCGLQIGRVMAINGIP 530
Qy 856 TEGMTHTQAVNLKNASGS-----TEMQVAGDVSVVTHHOBPASSLSFTGLTSTS 909
Db 531 TEDSTFEENQOLRDSITSKYTLEIEFDVAESVIPSSGTFHVKLPRKHSVEL-GITISS 589
Qy 910 IFODLGRP-----OCK--- 921
Db 922 -----SITLER--GPDGLGFSTVGGSPHGDLPYVTVFAK 957
Db 650 KIKIKEDNSDEQSSCAIITYVELKRGGLGITIS-----GTEPPDPPIISSLTG 704
Qy 958 GAASEDGLRKGDQIIAVNGQSLGEGVTHEEVAVALIKRTKGTVMV 1003
Db 705 GAERTGAIHIGRILAINSSSLKGPLSEDIHLLQMAGETVTLKI 750

```

RESULT 8  
 ; Sequence 15, Application us/09045632  
 ; Patent No. 6001575  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Hugenit, Richard L.  
 ; APPLICANT: Dong, Hualing  
 ; TITLE OF INVENTION: THERAPEUTIC USES OF GRIP AND  
 ; NUMBER OF SEQUENCES: 105  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
 ; STREET: 130 Water Street  
 ; CITY: Boston  
 ; STATE: MA  
 ; COUNTRY: USA  
 ; ZIP: 02109  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patentin Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/045.632  
 ; FILING DATE: 19-MAR-1998  
 ; CLASSIFICATION: 435  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 60/041.016  
 ; FILING DATE: 19-MAR-1997  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Corleiss, Peter F.  
 ; REGISTRATION NUMBER: 33,860  
 ; REFERENCE/DOCKET NUMBER: 48147/1699-CIP  
 ; TELECOMMUNICATION INFORMATION:

```

; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 702 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-045-632-15

```

Query Match 6.6%; Score 335; DB 3; Length 702;  
 Best Local Similarity 22.5%; Pred No. 1.9e-20;  
 Matches 171; Conservative 90; Mismatches 261; Indels 238; Gaps 22;

```

Qy 419 LELPRDQ-GIGIAISEEDTLGVI-----IKSLTEHVATDGRKLVGDQIIAVNDE 470
Db 3 VELMKKEGTTLC-----TVSGGIDKDGKPRVSNLRGGIARSDQDLVDGYIKAVNGI 56
Qy 471 IYVGPIREKFTLSLTKAKTVKLTTHAENPDQAVPSAAGAAGEKKNSQSLAMPQSGS 530
Db 57 NLAERHDEIISLLNNGERVLEVEYELP----- 86
Qy 531 PEPEIRNRSRSTPAIFASDPATCPRIIPGCTTIEISKRG--LGLSIVGSD---TL 585
Db 87 --PVSTQSS-----VME-----RTVEVTLHKGNTPGEVYIRGGAHDRNK 125
Qy 586 LGAFITIEYEEGAAKCDRLWAGDOILEVNGIDLRKATHDEAIWLRTPORVRLTL-Y 644
Db 126 SRPVYITCRPGPDRETIKRGDRLLSVDIGRLITTHAEMSLKCGGEATLIIY 185
Qy 645 RDEAPRKEEVEDTLTIELQKRGKGLSIVK--RNDTGVVSDIVYKGIADPDGRIL 702
Db 186 DVSADSVATASGPLVEAKTPGASLVALTTSVCCNKQVIVIDIKASASADRGALH 245
Qy 703 OGDQIILVNGEDVRNASQEAVALLKCSLGTVLEY-----GRK-AGPFH--- 747
Db 246 VGDHILSIDGTSMETCTLAETQFLGNTDQVKLELPHHQRLALKGDDHVKIORSDQ 305
Qy 748 -----SERPSQSOYSE 760
Db 306 LPMDPWASQCSVHTNNHHNPHHCHRVALGFPKALTPNSPRAWSSSPTSMSAYSL 365
Qy 761 GSLSTFPPLSGSSTS-----ESLESSKKNAASEIOLKRV--MKK 802
Db 366 SSLNMGTLPRSLSTSPRCTMMRRLLKKDFKSSLSLASTVGLAQVHTETTEVLTA 425
Qy 803 GPTDSIGISAGV--GSPLDGVPIFTAMHPTGVAOQOKLVRGDRIVTICSTEGMTH 861
Db 426 DRYTGIGIOLQSGVFATETLSFPLISYLEADSPARCGLQIGRVMAINGIPEDSTF 485
Qy 862 TQAVNLKNASGS-----TEMQVAGDVSVVTHHOBPASSLSFTGLTSTSIFODL 915
Db 486 EKANOLLRDSITSKYTLEIEFDVAESVIPSSGTFHVKLPRKHSVEL-GITISSPSRKP 544
Qy 916 GPP-----OCK--- 921
Db 545 GGPLVVISDIKKSVAHRTGTLLEGDLAIDNIRLDCSMEDAVOILQCEDELVKIKRK 604
Qy 922 -----SITLER--GPDGLGFSTVGGSPHGDLPYVTVFAKGAASED 963
Db 605 DEDNSDEQSSCAIITYVELKRGGLGITIS-----GTEPPDPPIISSLTGKGLAENT 659
Qy 964 GLRKGDQIIAVNGQSLGEGVTHEEVAVALIKRTKGTVMV 1003
Db 660 GAHIGRILAINSSSLKGPLSEDIHLLQMAGETVTLKI 699

```

RESULT 9  
 ; Sequence 16, Application us/09045632  
 ; Patent No. 6001575

```

: GENERAL INFORMATION:
: APPLICANT: Huguair, Richard L.
: TITLE OF INVENTION: THERAPEUTIC USES OF GRIP AND
: NUMBER OF SEQUENCES: 105
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
: STREET: 130 Water Street
: CITY: Boston
: STATE: MA
: COUNTRY: USA
: ZIP: 02109
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/045,632
: FILING DATE: 19-MAR-1998
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 60/041,016
: FILING DATE: 19-MAR-1997
: ATTORNEY/AGENT INFORMATION:
: NAME: Corless, Peter F.
: REGISTRATION NUMBER: 33,860
: REFERENCE/DOCKET NUMBER: 48147/1699-CIP
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 617-523-3400
: TELEFAX: 617-523-6440
: INFORMATION FOR SEQ ID NO: 16:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1018 amino acids
: TYPE: amino acid
: STRANDEDNESS:
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-09-045-632-16

```

```

Query Match      6.6%; Score 335; DB 3; Length 1018;
Best Local Similarity 22.5%; Pred. No. 3.6e-20;
Matches 171; Conservative 90; Mismatches 261; Indels 238; Gaps 22;

```

```

QY 419 LELPKDGG-GGIAIASEDDTSGYI-----IKSLTEHGVAATDGRKAVDQDLAVDDE 470
DB 3 VELKKKRETTLGC-----TVSSGIDKDGKPRVSNLRQGGIAASDDLDVGDYIKAVNGI 56
QY 471 IVVGYPIEKFTSLTKAKMTVKLTIIHAENPDQAVPSNAGAAGEKKNSSQSLWVPOGS 530
DB 57 NLAQFRDEIISLKNNGERYVLEVEYELP----- 86
QY 531 PEPISIRNTSSSPRAIPASDPATCPIIPGCEETIETISKGTG--LGISTVGGSD--TL 585
DB 87 --PVSIGSS-----VNF-----RIVEVTLHKEGTGFVYIGGAHDRNK 125
QY 586 LGATIIHEVVEEAGACDGRIMAGDQILEVNGIDLRKATHEALINVLKQTPQRRRLTL-Y 644
DB 126 SRPVYITCVPRGGPDDEGTTPKPGDRILSVGIRLLGTTHAEANSLIKQCEQETLLEY 185
QY 645 RDEAPYKEEVEDTLTTELQKPKGKGLSLVGR--RNDTGVFSDIYKGIADPDGRILI 702
DB 186 DVSAMDVATASGRLVLEVAKTPGASIGVATITVSCNKQIVYDKIKSASIAIRCALH 245
QY 703 QGDQIILVNGEDVARNASQEAVALLLKCSLGTVTLEV-----GRIK-AGPRH----- 747
DB 246 VGDHILSIDGSMETCYLAETQPLGNTTPOVKLEILPHHQTRIALKGPDHVKIQRSDRQ 305
QY 748 -----SERPSTSOVSE 760
DB 306 LPMDPMASSQCSVTHNHNHNPDRHCHRVPALEGPKALTPNSPAMVSSSPSTMSAYSL 365

```

```

QY 761 GLSISFTEPLSGSSNS-----ESLESSKKNALASEIQLRTVE--MKK 802
DB 366 SLMNGITPRSLYSTSPGTYMMRRRLKKDKRSLSLASISVGLAGOVYVETTEVYILA 425
QY 803 GPTDSLGISIAGV-GSPPLGDPPIFIAMHPTGVAQAOTKLRVDRIVYITIGSTEGMTH 861
DB 426 DPVGVGIGLOGSVFATETLSSPLISYIEADSPAERCGVLIQIDRVVAINGIPTEDSTF 485
QY 862 TQAVNLKNAGS-----IMQYVAGDVSVYGHNOEPASSLSIFGLSTISIFDDL 915
DB 486 EEAQQLLRDSSITSKYTLEIFDVAESVYIPSSGTFHVKLPKRHSVEL-GTISIPSSSRKP 544
QY 916 GPP-----QCK----- 921
DB 545 GDPPLVSDIKKGSVAHRTGTLELGDKLAIIDNRLDSCSMEDVAQIILQCCEDLYKIKRK 604
QY 922 -----SITLER--GPDGLGFSIVGYGSPHGDLPYVKTVFAKGAASED 963
DB 605 DEBNSDEQESSGAILIYVELKRYGPGIDITIS-----GTEEPFPIIISLTKGGLAERT 659
QY 964 GRLKRGDQIIAVNGQSLGCVTHEEVAVALKRTKGTVTLMV 1003
DB 660 GAHIGDRILAINSSSLKGRPLSEDIHLQAGETVYTKI 699

```

```

RESULT 10
US-09-045-632-32
: Sequence 32. Application US/09045632
: Patent No. 6001575
: GENERAL INFORMATION:
: APPLICANT: Huguair, Richard L.
: TITLE OF INVENTION: THERAPEUTIC USES OF GRIP AND
: NUMBER OF SEQUENCES: 105
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
: STREET: 130 Water Street
: CITY: Boston
: STATE: MA
: COUNTRY: USA
: ZIP: 02109
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: IBM PC compatible
: SOFTWARE: Patentin Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/045,632
: FILING DATE: 19-MAR-1998
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 60/041,016
: FILING DATE: 19-MAR-1997
: ATTORNEY/AGENT INFORMATION:
: NAME: Corless, Peter F.
: REGISTRATION NUMBER: 33,860
: REFERENCE/DOCKET NUMBER: 48147/1699-CIP
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 617-523-3400
: TELEFAX: 617-523-6440
: INFORMATION FOR SEQ ID NO: 32:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1061 amino acids
: TYPE: amino acid
: STRANDEDNESS:
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-09-045-632-32

```

```

Query Match      6.6%; Score 335; DB 3; Length 1061;

```



Best Local Similarity 22.5%; Pred. No. 3.9e-20;  
Matches 171; Conservative 90; Mismatches 261; Indels 238; Gaps 22;

|    |     |  |      |
|----|-----|--|------|
| 0Y | 419 | LELPKODG-GGAIMISEBDTSLGYI-----IKSLTSEHVAAATDGLKJYDQOILLVDE     | 470  |
| Db | 3   | VELKKREBTLLGC-----IYSGGIDDKGPRVSNLRQGGIAARSDLDVGYITKAVNGI      | 56   |
| 0Y | 471 | IVGVPLEKFIISLLTKAKMTYVKTLTHAENPDSQAVPSAAGAAGEKKNSQSILMPQSGS    | 530  |
| Db | 57  | NLAFFRHEDEIISLLKNVGERVULEVEYELP-----                           | 86   |
| 0Y | 531 | PEPESIRNTSSSRPALFPADSPATCPIITPCETTIEISKRTG--LGLSVGSD---TL      | 585  |
| Db | 87  | --PVSIGGS-----VWF-----RTVEVTLTKKEGTGEFVIRGAHDHRNK              | 125  |
| 0Y | 586 | LGAFIHEHYEGEAACDGRRLMAGDOILLENGIDLRKATRDEAIINVLRTPQGVRLTL-Y    | 644  |
| Db | 126 | SRPVYITCYVRGGRDDREGTIKPBDRLISVDGITLGTTHAEMSLIKQCGQCATLLIEY     | 185  |
| 0Y | 645 | RDEAPRYKEEVEDTLTLELOKPKGKGLSLGYK--RNDTGVFVSDIVKGGIADPDGRILI    | 702  |
| Db | 186 | DVSAAMDVSATASGFLVEAVATPGASLGVALLTYSVCCNKQVYIDIKSAISADRCGALH    | 245  |
| 0Y | 703 | QGDQILLVNGEDVYRNASOEAVALLKCSLGTVILEY-----GRIK-AGPRH-----       | 747  |
| Db | 246 | VGDHILSIDGTSMEYCTLAETOPFIGNTTQVUKLEILPHHQTRIALKGPDHVHTIORSRQ   | 305  |
| 0Y | 748 | -----SERPKOTQSVSE  | 760  |
| Db | 306 | LPMDPMASSQCSYHTNNHHNPHRPHDCHRPALGPRLKTLPNSPRPMWSSSPTSMSAYSL    | 365  |
| 0Y | 761 | GSLSSEFTPLSGSSTS-----ESLESSKKNMLASEIQLGRIVE--MKK               | 802  |
| Db | 366 | SSLNMGTLPRSLYSTSPRGTMRRRLKKPKDKSSLASSPYGLAGVYHTEETVELTRA       | 425  |
| 0Y | 803 | GPTBSLGISTINGVY-GSPFLGDVPRFFIMMHPRTGYAAOTQKLRVGDRLVYTCGSTBGMTH | 861  |
| Db | 426 | DPVYFGFIQLOGSVFAFETLSSPLISYIEADSPAEKGVLOJDRYMAINGIPTEDSTF      | 485  |
| 0Y | 862 | TOAVNLKKNAGGS-----IEMQVVAAGDVSVYVGHQOEPAASSLSLSTGLTISFODDL     | 915  |
| Db | 486 | EEAQNOLDRDSSITSKVYLIEFDVAESVYFSSGFIHKLKLRKKSVEL-GITTISSPSSRKP  | 544  |
| 0Y | 916 | GPP-----OQR-----   | 921  |
| Db | 545 | GDPVLVTSIDIKKGSVAHRTGTELEDBKLKLAIDNRKLDSCSMEDAVQILOQCEDLVKLIRK | 604  |
| 0Y | 922 | -----STTLER--GPRGLGFSYVGSGSPHGDLPIYVTKVFPKGAASED               | 963  |
| Db | 605 | DEDSNDEQESSGAIIVYELKRYGGRPLGITIS-----GTEEPDPPIITISLTKGLAERT    | 659  |
| 0Y | 964 | GLRLKRGDOIIIVANGOSLEGVTHEAVANMLKRTKGTVILMY                     | 1003 |
| Db | 660 | GAHIGDRITLAINSSSLKGRPLSEDIHLLQVAGCTVYLUKI                      | 699  |

RESULT 11  
 US-08-410-804-1  
 : Sequence 1, Application US/08410804  
 : Patent No. 5632994  
 :  
 : GENERAL INFORMATION:  
 : APPLICANT: Reed, John C.  
 : APPLICANT: Sato, Takaki  
 : TITLE OF INVENTION: FAS ASSOCIATED PROTEINS  
 :  
 : NUMBER OF SEQUENCES: 22  
 :  
 : CORRESPONDENCE ADDRESS:  
 : ADDRESSEE: Cathryn Campbell  
 : STREET: 4370 La Jolla Village Drive, Ste 7000  
 : CITY: San Diego  
 : STATE: California  
 : COUNTRY: United States  
 : ZIP: 92122  
 :  
 : COMPUTER READABLE FORM:

```

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/410,804
FILING DATE: 27-MAR-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/259,514
FILING DATE: 14-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-LJ 1389
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 610 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-410-804-1

```

|                            |       |                   |             |            |
|----------------------------|-------|-------------------|-------------|------------|
| Query Match                | 6.6%  | Score 334.5       | DB 1        | Length 610 |
| Best Local Similarly       | 23.0% | Pred. No. 1.7e-20 |             |            |
| -Matches 155; Conservative | 72;   | Mismatches 221;   | Indels 225; | Gaps 21.   |

|    |     |   |      |
|----|-----|---|------|
| QY | 474 | GYPTEKFEISLTKTAKMYVKKLIIHAENPDSQAVPAMAGAAGEKK-----NSSQSL      | 5333 |
| Db | 2   | GSPSPSVIS-----KATEKEFTFDSNGSKTKKPGISDVDTYSDRGSDMDQATYSSQDH    | 56   |
| QY | 524 | MYPOSGSEPEPSINTSKSSSTPAIFADDPACPIIPCCETTHIELSKRTGLGSLIV----   | 579  |
| Db | 57  | QTPKO-----ESSSVNTSKMKNKFTSSPPK-----PDIFEEVELAKDNLSIGISVTLFD   | 1099 |
| QY | 580 | -GGSDDL--GAFTHIEYEEGAACKGGRRLMAGDOILLEVGIDLRKATPHDEAIVNLTOTP  | 6366 |
| Db | 110 | KGVNTSVIRHGGIYVKAIVPGGAESDGRJHKGDRLVLANVGSLEGATTHQAVETLNTG    | 1659 |
| QY | 637 | QVRLITLYRDAFPRKEEV-----C-----DLTI                             | 661  |
| Db | 170 | QVHLLLEKGGSPTSKEHVPTPOCTLSDONAGQGEKYAKKTQVXKDYSPVTEENFEV      | 229  |
| QY | 662 | ELOKKPGKGLS-----IVGRNDTGVPSIVYGGIADPDGRJLISGOLILVNGED         | 714  |
| Db | 230 | KLFKN--SSGLCFSTSRDNILPEIDNMSYAVKRLFPGQPAESGKIDVDVILKVGAS      | 2888 |
| QY | 715 | VRNASEAVALLKCSLGTVTLLEVGRKAGPF-----HSERRSQT                   | 7555 |
| Db | 289 | LKGSQOEIVATLKGTAPEVFLLCRPPGVLPEIDTALLTPLOSQAQVLPNSKDSOSP      | 3488 |
| QY | 756 | SOVEBGLSSTFFLGSSTSEESLESSKKALLASELOGLRTVMKGPDSLGISIAG         | 815  |
| Db | 349 | SCVED-----STSDENENSDSKKQC-----KSPSRDYSOSSG                    | 3844 |
| QY | 816 | VGSPGADVPIFIAMHPTGVAQAQTKRLVCGRIYITIGSTEGMTHQAVNLKNAAGSI      | 8757 |
| Db | 385 | SGE-----DDLVY-----APANISNWTWS                                 | 4004 |
| QY | 876 | EMOVYAGDVSVTYVGHQOEPASSLSSTGL-----TSTFSIPODDLGP--             | 917  |
| Db | 405 | ALHQTLSNMVSOAQSHHDPKRSQEDTICTMYEYPOKIPNKRPEFDSNPSPLPMPMAGQS   | 4644 |
| QY | 918 | --POCKS-----ITLERPDG  | 931  |
| Db | 465 | XQPOSEGASSSMOKYHHTHSTPTROENMPLKNDLENHLEDEBELEVELLITLIKESG     | 5344 |
| QY | 932 | -LGFISVGGYGSPIHGDLPIYKVTYFAKGAASEDGRKRGDOITIAVNGOSLEGYTHEEAVA | 9900 |

Db 525 SLGFTVTKG---NORIGCYHVDI-QDPAKSDGRKRGRLIKVNDVTNMTHTDAVN 579  
QY 991 ILKRTKGTVTLMV 1003  
Db 580 LLRAASKTVRLVI 592

RESULT 12  
US-08-259-514-1  
Sequence 1, Application US/08259514  
Patent No. 5747245  
GENERAL INFORMATION:  
APPLICANT: Reed, John C.  
APPLICANT: Sato, Takaaki  
TITLE OF INVENTION: FAS ASSOCIATED PROTEINS  
NUMBER OF SEQUENCES: 22  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Cathryn Campbell  
STREET: 4370 La Jolla Village Drive, Ste 700  
CITY: San Diego  
STATE: California  
COUNTRY: United States  
ZIP: 92122  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/259,514  
FILING DATE: 14-JUN-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Campbell, Cathryn  
REGISTRATION NUMBER: 31,815  
REFERENCE/DOCKET NUMBER: P-LJ 9954  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 535-9001  
TELEFAX: (619) 535-8949  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 610 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-259-514-1

Query Match 6.68; Score 334.5; DB 1; Length 610;  
Best Local Similarity 23.08; Pred. No. 1.7e-20;  
Matches 155; Conservative 72; Mismatches 221; Indels 225; Gaps 21;

QY 474 GYPIEKFTSLTKAKMTVKLTIIHAENPDQAVPAAGAAGEK-----NSSQSL 523  
Db 2 GSPSPVIS-----KATEKETFTDSNQSRTKPKISDVTDVSDGSDMDATYSSSDH 56  
QY 524 MYPOSGSPESINSTRSSSTPAIFASDPATCPIIPGCEETIEISKGTGLGISIV---- 579  
Db 57 QTPKQ---ESSSYNTSNKNMFKTFSSSPK---PGDIFEVELAKNDNSLGISVTVLFD 109  
QY 580 -GGSDTL--GAFIIHVEYEGACKDGRLMAGDQILLEVNGIDLRKATHDAINVLQTP 636  
Db 110 KGVNTSVRRHGIVYKAVIPGGAESDGRJHGDVLAIVNSLEGATHKQAVETLRNTG 169  
QY 637 QRVALTLYRDAPYKEEV-----C-----DLTI 661  
Db 170 QVAVLLLEKGGSPTRKEHVPTTQCTILSDONAOGQPEKVKTTQVNDYSTVEENTFEV 229  
QY 662 ELQKKPKGJGLS-----IVGRNDTVGVFVSDIVKGLADPDGRLIQDQILLVNGED 714  
Db 230 KLFKN-SSGIGFSRSDNLIPEQINASIYVKKLPGQPAESGKIDVGVILKVGAS 288  
QY 715 VRNASQAVALLKCSIGTIVLEVGRIKAGF-----HSRRPSQT 755

Db 289 LKGLSQDEVISALRGTAEVLLLCRPPGVLPEDTALLTPLOSPAQVLPNSSKSSQP 348  
QY 756 SOVSEGLSFTPLSGSTSESSSKKNANLASEIGLRTVEKKKQPTDSLGISING 815  
Db 349 SCVED-----STSDENEMSKSKOC-----KSPSRDSTSDSG 384  
QY 816 VSPPLGDVPIFIAMHPTGVAAGTOKLRVGDRIYICGTSTEGMTHQAVNLLNAGSI 875  
Db 385 SGE-----DDLVY-----APANISSTWSS 404  
QY 876 EMQVAGDVSVTGHQEPASSLSFTGL-----TSTSTQDGLP--- 917  
Db 405 ALHQTLSNMVSAQASHREAPKSOEDTICTMYYPQKIPNKEPFEEDSNPSLPDPAPOS 464  
QY 918 --POCKS-----ITLERGPDG 931  
Db 465 YQPOSESASSSMKRYHIIHISEPTROENWTPLNKDLENHLEDELEVELLITLIKSEK 524  
QY 932 -LGFSTVGYGSPHDDLPYVYKFAKGAASEDRKRGDQIIAVNGSLGVTHEEAVA 990  
Db 525 SLGFTVTKG---NORIGCYHVDI-QDPAKSDGRKRGRLIKVNDVTNMTHTDAVN 579  
QY 991 ILKRTKGTVTLMV 1003  
Db 580 LLRAASKTVRLVI 592

RESULT 13  
US-08-858-311-1  
Sequence 1, Application US/08858311  
Patent No. 5876939  
GENERAL INFORMATION:  
APPLICANT: Reed, John C.  
APPLICANT: Sato, Takaaki  
TITLE OF INVENTION: FAS ASSOCIATED PROTEINS  
NUMBER OF SEQUENCES: 22  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Cathryn Campbell  
STREET: 4370 La Jolla Village Drive, Ste 700  
CITY: San Diego  
STATE: California  
COUNTRY: United States  
ZIP: 92122  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/858,311  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Campbell, Cathryn  
REGISTRATION NUMBER: 31,815  
REFERENCE/DOCKET NUMBER: P-LJ 1389  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 535-9001  
TELEFAX: (619) 535-8949  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 610 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-858-311-1

Query Match 6.6%; Score 334.5; DB 2; Length 610;  
 Best Local Similarity 23.0%; Pred. No. 1.7e-20;  
 Matches 155; Conservative 72; Mismatches 221; Indels 225; Gaps 21;

QY 474 GYPIKFTSLTAKTAVLTTHAENPDQAPSAAGASGEK-----SSQSL 523  
 DB 2 GSPSPSVIS-----KATEKETFTDSNQSKTKKPGISDVTDYSDRQSDMDKDEATYSSQDH 56  
 QY 524 MYPOGSEPESTRNSTRSPAIRASDPATPIIPGCTTEISKGTGLSLV---- 579  
 DB 57 QTPKQ---ESSSVTSMKMFTESSPPK---PGDJFEVELAKNDMSLSIVTVLED 109  
 QY 580 -GGSTFL--GAFTIHEVEEGACKDGLMAGDOILEVNGIDLKATHEAIVNLRORP 636  
 DB 110 KGVNTSVAHGGITVYKAVIPQGAESDGRTHKGRVLAIVGSLGATHKQAVETLRNG 169  
 QY 637 QVRRLTLXDEAPYKEEVY-----C-----DLETI 661  
 DB 170 QVYHLLLEKGSPTSKENHVPVTPQCTLSQDQNAQSGPEKKTQYKQDSFVTEENTFEV 229  
 QY 662 ELQKRPKGLIS-----IVKRNQGVYSDIYKGIADPDGRLQDQILLVNGED 714  
 DB 230 KLFKN-SSGLGFSFREDNLIPEQINASTYRYKKLFPQGPAAESGKIDVDYLKNGAS 288  
 QY 715 VNASQEAVALKCSLGTVTLEVGRIKAGPF-----HSRRPSQT 755  
 DB 289 LKGLSQQEVYSLKGTAPVFLILCHPPGVLPETIDTALLTLQSPQVLPNSSKDSQP 348  
 QY 756 SOVSEGLSLTFPPLSGSSTSESSSKKNALASEIQLRTVEKKGPTDLSIGISAG 815  
 DB 349 SCVEQ-----STSDENEMSDSKSKQC-----KSPSRDYSDSG 384  
 QY 816 VSPPLGDVPIFTAMHPTVANAQOTKLRGSDHIVTICGSTEGMTHTQAVNLLKNSGI 875  
 DB 385 SE-----DLVT-----APANISNTWSS 404  
 QY 876 EMQVAGDVSVYVTHQBPASSLSFTGL-----TSTIFODDLR--- 917  
 DB 405 ALHQTLNNVNSQAQSHENAKSQEDTICMFTYYPQKIRPKPEREDSNPRLPPDPAQGS 464  
 QY 918 --PQCKS-----ITLERGPDG 931  
 DB 465 YQPOSASASSSMKYHHIHPSEPTROENMTPLKNDLENHLEDFELEVELLTLIKSEK 524  
 QY 932 -LGFSTVYGSPHBDLPYVKTVPKGAASEGRKRDQIIIVAGSLGEGTHEANA 990  
 DB 525 SLGFYTKG---NORIGCYVDVI-QDPAKSDGRLKPDRLKIVNDTIVNTHTDAVN 579  
 QY 991 ILKRTGVTLMV 1003  
 DB 580 LRAASKTVRLVI 592

RESULT 14  
 US-09-045-632-50  
 Sequence 50, Application US/09045632  
 Patent No. 6001575  
 GENERAL INFORMATION:  
 APPLICANT: HUGANIR, RICHARD L.  
 APPLICANT: DONG, HUALING  
 TITLE OF INVENTION: THERAPEUTIC USES OF GRIP AND  
 TITLE OF INVENTION: GRIP-RELATED MOLECULES  
 NUMBER OF SEQUENCES: 105  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
 STREET: 130 Water Street  
 CITY: Boston  
 STATE: MA  
 COUNTRY: USA  
 ZIP: 02109  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/045.632  
 FILING DATE: 19-MAR-1998  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 60/041,016  
 FILING DATE: 19-MAR-1997  
 ATTORNEY/AGENT INFORMATION:  
 NAME: CORLESS, Peter F.  
 REGISTRATION NUMBER: 33,860  
 REFERENCE/DOCKET NUMBER: 48147/1699-CIP  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 617-523-3400  
 TELEFAX: 617-523-6440  
 INFORMATION FOR SEQ ID NO: 50:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1050 amino acids  
 TYPE: amino acid  
 STRANDEDNESS:  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-09-045-632-50

Query Match 6.6%; Score 334.5; DB 3; Length 1050;  
 Best Local Similarity 21.9%; Pred. No. 4.2e-20;  
 Matches 203; Conservative 129; Mismatches 321; Indels 275; Gaps 42;

QY 157 VELMREPSLSISIVGR-----GMGR---LSNGEVMR---GIFKHYL 196  
 DB 89 VELIKREGSTLGTISGIDTKDKKPRVSNLRREGGLARSDLVNGDIYSVNGIRLTRR 148  
 QY 197 EDSPPAKNGTL--KPGDRIVEAPSOSESSEPERKAPLCSVPPPSAPFAENGSDHTQSSASK 254  
 DB 149 HBEII---TLKNVGERVY---LEVEYE---LPPAP-----ENNPI 182  
 QY 255 ISQDVKDEDFGYSKNIHERYGTGLGELHMELEKGGHGLSLAG---NKDRRMV 310  
 DB 183 ISKTYD-----VSLKEGNSFGVGLKGAHEDLHKSRPLV 217  
 QY 311 FYIGIDPNCAGAGDGRLOADELLEINGQILYGRSHQNA--SLIKCAPSKVKIIFRNKD 369  
 DB 218 -LTYVAPGPANREGSLKGDRLSLXGIPRHGASHATATATLQCC-----SHE 265  
 QY 370 AVNQAVPCGNAVEPLPSNSENLOKNETEPTVYTTSDAAYDLASFKNVQHLLEPKDGGGLG 429  
 DB 266 ALFOY-----EYDATPPTVANASGPLVEIATKPGSALGIS 302  
 QY 430 IAISEDTLSGYIILKSLTEHGAATDGRKLVGDQIIIVADDEIVYGPICEKFTSLTKAM 489  
 DB 303 LFTGSHRNKPAITIDIRIKPASVVDKNGALHAGEHIIAIDGTSTHESLVEATKLASVTE 362  
 QY 490 TYKLTTHAENPDQ---AVPSAAGASGEKKNSSQSLAMYPOGSEPESTR--- 537  
 DB 363 KVRLEI-LPAQSRKRLKPEAVRIQSRQDLHMDP-CVPSCHSPRSHCRAPTAPGGO 420  
 QY 538 -NTSRSPATFAFSDPATPIIPGCTTEISKG---RT----- 572  
 DB 421 DQSRVSTP---FSSPTNNPAP-CANASTLDRGMSRRTTAGRRORRKHRSLSLA 476  
 QY 573 -----GLGLSTYVG---SDTLGAFITHIEVEEGACAD 603  
 DB 477 SSTVPGGOIVHTEETEVVLCGDPISGFGIQLQGGIFATETLSSPLVFRIBDSPAEC 536  
 QY 604 GRMLAGDOILEVNGIDLKATHEAIVNLRQ---PQVRRLTLXDEAPYKEEVY---CDT 658  
 DB 537 GLQVGDRLVLAINGIATEGCTHEANQILRDAALAKIYLEIFDYA-----BSVIPSSGT 592  
 QY 659 LTELQKRPKGLISL--VGRNDTGVFVSDIVKGIADPDGRLQDQILLVNGEDVR 716

```

Db 593 FHVKLPKRGEVLEGTITSSASRKKEPILLISDIKKSVAAHRTGLEPDKLAIIDNIRLD 652
QY 717 NASOE-AVAALLKCSLGTVTLEVGRIKAGPHSRERSQTS--QVSEGSLSFTFPLSGS 773
Db 653 HCPMEYAVOILPOCE-DLVKLI-----RKDEDNDEOESSGAVS-----691
QY 774 STSELESSSKNNALASEIOGLRTVEMKK--GPTDSLGISIAGVSPGLGDPFIAMH 831
Db 692 -----YVELKRYGCP--LGITIS-GTEEPFD--PIIISGLT 723
QY 832 PTGVAOQOKLRVGDRIYVTCGTSTEGMTHQAVNLKNAGSITEMOV-----A 881
Db 724 KRGLAERTGASMLGTAWPSXSVSLKGRPLSEAIHLQVAGETVTLKIKQIDRPLLRQ 783
QY 882 GADVSVVTHGHEBPASSLSFTGLTSTIFQDDIGPPOCKSITLERGPDGLGSIV-GGY 940
Db 784 SGSLSEASDVDEDPPEALKG--GLTTTHF-----SPAVSV--DSAVESMGSSATEGCF 833
QY 941 GSPHGDLPYVKYTFAPKGAASEDRGLKR 968
Db 834 GSGS-----YTPQVAVRSVTPQEWRRSR 857

RESULT 15
US-09-045-632-49
: Sequence 49, Application US/09045632
: Patent No. 6001575
: GENERAL INFORMATION:
: APPLICANT: Hugent, Richard L.
: APPLICANT: Dong, Hualing
: TITLE OF INVENTION: THERAPEUTIC USES OF GRIP AND
: TITLE OF INVENTION: GRIP-RELATED MOLECULES
: NUMBER OF SEQUENCES: 105
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
: STREET: 130 Water Street
: CITY: Boston
: STATE: MA
: COUNTRY: USA
: ZIP: 02109
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent in Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/045,632
: FILING DATE: 19-MAR-1998
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 60/041,016
: FILING DATE: 19-MAR-1997
: ATTORNEY/AGENT INFORMATION:
: NAME: Corleese, Peter F.
: REGISTRATION NUMBER: 33,860
: REFERENCE/DOCKET NUMBER: 48147/1699-CIP
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 617-523-3400
: TELEFAX: 617-523-6440
: INFORMATION FOR SEQ ID NO: 49:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1050 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-09-045-632-49

```

```

Query Match 6.5%; Score 331.5; DB 3; Length 1050;
Best Local Similarity 21.8%; Pred. No. 7,7e-20;
Matches 202; Conservative 130; Mismatches 321; Indels 275; Gaps 42;
QY 157 VELMREPSKSLGISIVGR-----GMGRS---LSNGEYMR---GIFIKHVL 196

```

```

Db 89 VELIKREGSTLGLTISGTDKDGKPRVSNLAPGGLAARSDLANVGDYIRSVNGIRLRLR 148
QY 197 EDSAGKNGTL--KPGRIYAPASQSEPEKAPLCSYPPPEPFAFEMGSDHTQSSASK 254
Db 149 HDEII-----TLKNVGERV-----LEVEYE-----LPPAP-----ENNRI 182
QY 255 ISOVDKEDERFGYSKMNIRERYGTLTGELHMELEKSHGSLAG-----NKDRSMV 310
Db 183 ISKTYD-----VSLYKGNSTGFVLRGAEHDLKSRPLY 217
QY 311 FIVGIDPNGAAGKDGRLQIADELLEINGQILYGRSHQNA--SIIKCAPSKYKIIIFIRNKD 369
Db 218 -LTYVRPGGPANREGSLKVGDRLLSIXGILPHGASHATAITLQOC-----SHE 265
QY 370 AVNMAVCPGNAVEPLPSNSNLQNKKEPEYVTSDAVDSLKKYQHLELPKDGGLG 429
Db 266 ALFOV-----EYDVATPDVANAAGPLVVEIATTPSALGIS 302
QY 430 IAISEEDTLGVIILKSLTEHGAATDGRLLKVGQIILAVDEIVVGYPIEKFTSILKTAKM 489
Db 303 LTTGSHNNKPAITIDRIKPAVYDKNALNGEHIILAIIDGTSTHCSLVEATKILASYTE 362
QY 490 TVKLTIAHENDSQ--AVPSAGAAGEKKNSOSLAVPOGSPPEPESIR-----537
Db 363 KVRLEI-LPAQSRRLKPEPAVRIORSEOLHMDP-CVPSCHSPRPSHCAPTMAPGQ 420
QY 538 --NTSRSTPAIFASDPATCPIIGCETIISG-----RT-----572
Db 421 DQSRVSTP--FSSPTMNAFP-CANASTLPKGPSPPTTACRRQRKRHRSSLSLA 476
QY 573 -----GLGLSIVG-----SDTLGAFIHEYEGAAKD 603
Db 477 SSTVPGQIYHTEVTEVLCGDPALSGQLQGIATETLSSPLVRFIEPSPAREC 536
QY 604 GRIMAGQIILEVNGIDLRKKAHDAIVLROT--PQVRKLTLYDEAPYKEEY--CDT 658
Db 537 GLLOVGDVLAINGIATEDGTMEQANOLLRPAALARKIVLEIEFDVA-----ESVYPSGT 592
QY 659 LTILOKPKGKGLSTI--VGRNDTGVEFVDIYKGIADPDGRLIGQDQIILVNGEDVR 716
Db 593 FHVKLPKRGEVLEGTITSSASRKKEPILLISDIKKSVAAHRTGLEPDKLAIIDNIRLD 652
QY 717 NASOE-AVAALLKCSLGTVTLEVGRIKAGPHSRERSQTS--QVSEGSLSFTFPLSGS 773
Db 653 HCPMEYAVOILPOCE-DLVKLI-----RKDEDNDEOESSGAVS-----691
QY 774 STSELESSSKNNALASEIOGLRTVEMKK--GPTDSLGISIAGVSPGLGDPFIAMH 831
Db 692 -----YVELKRYGCP--LGITIS-GTEEPFD--PIIISGLT 723
QY 832 PTGVAOQOKLRVGDRIYVTCGTSTEGMTHQAVNLKNAGSITEMOV-----A 881
Db 724 KRGLAERTGASMLGTAWPSXSVSLKGRPLSEAIHLQVAGETVTLKIKQIDRPLLRQ 783
QY 882 GADVSVVTHGHEBPASSLSFTGLTSTIFQDDIGPPOCKSITLERGPDGLGSIV-GGY 940
Db 784 SGSLSEASDVDEDPPEALKG--GLTTTHF-----SPAVSV--DSAVESMGSSATEGCF 833
QY 941 GSPHGDLPYVKYTFAPKGAASEDRGLKR 968
Db 834 GSGS-----YTPQVAVRSVTPQEWRRSR 857

```

Search completed: July 12, 2001, 14:41:11  
Job time: 214 sec

Fri Jul 13 15:00:11 2001

us-09-502-698-2.ra1

Page 13



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 12, 2001, 18:54:42 ; Search time 6307.92 Seconds

(without alignments)  
11966.324 Million cell updates/sec

Title: US-09-502-698-3

Perfect score: 4880

Sequence: 1 cccggcgccggcgagactg.....tctctcccccacccaact 4880

Scoring table:

IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1344157 seqs, 7733874588 residues

Total number of hits satisfying chosen parameters: 2688314

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:\*

1: gb\_ba1:\*  
2: gb\_ba2:\*  
3: gb\_ba3:\*  
4: gb\_in1:\*  
5: gb\_in2:\*  
6: gb\_in3:\*  
7: gb\_om:\*  
8: gb\_ov:\*  
9: gb\_pat1:\*  
10: gb\_pat2:\*  
11: gb\_ph:\*  
12: gb\_pl1:\*  
13: gb\_pl2:\*  
14: gb\_pl3:\*  
15: gb\_pl4:\*  
16: em\_ba1:\*  
17: em\_ba2:\*  
18: em\_fun:\*  
19: em\_htgo\_hum:\*  
20: em\_htgo\_inv:\*  
21: em\_htgo\_rod:\*  
22: em\_htg\_hum1:\*  
23: em\_htg\_hum2:\*  
24: em\_htg\_hum3:\*  
25: em\_htg\_hum4:\*  
26: em\_htg\_hum5:\*  
27: em\_htg\_hum6:\*  
28: em\_htg\_hum7:\*  
29: em\_htg\_hum8:\*  
30: em\_htg\_inv1:\*  
31: em\_htg\_inv2:\*  
32: em\_htg\_other:\*  
33: em\_htg\_rod:\*  
34: em\_hum1:\*  
35: em\_hum2:\*  
36: em\_hum3:\*  
37: em\_hum4:\*  
38: em\_hum5:\*  
39: em\_hum6:\*  
40: em\_hum7:\*  
41: em\_in:\*  
42: em\_om:\*  
43: em\_or:\*

44: em\_ov:\*  
45: em\_pat:\*  
46: em\_ph:\*  
47: em\_pl:\*  
48: em\_ro:\*  
49: em\_sts:\*  
50: em\_sy:\*  
51: em\_un:\*  
52: em\_v1:\*  
53: gb\_sts1:\*  
54: gb\_sts2:\*  
55: gb\_sts3:\*  
56: gb\_sy:\*  
57: gb\_un:\*  
58: gb\_v11:\*  
59: gb\_v12:\*  
60: gb\_htg1:\*  
61: gb\_htg2:\*  
62: gb\_htg3:\*  
63: gb\_htg4:\*  
64: gb\_htg5:\*  
65: gb\_htg6:\*  
66: gb\_htg7:\*  
67: gb\_htg8:\*  
68: gb\_htg9:\*  
69: gb\_htg10:\*  
70: gb\_htg11:\*  
71: gb\_htg12:\*  
72: gb\_htg13:\*  
73: gb\_htg14:\*  
74: gb\_htg15:\*  
75: gb\_htg16:\*  
76: gb\_htg17:\*  
77: gb\_htg18:\*  
78: gb\_htg19:\*  
79: gb\_htg20:\*  
80: gb\_htg21:\*  
81: gb\_htg22:\*  
82: gb\_htg23:\*  
83: gb\_htg24:\*  
84: gb\_htg25:\*  
85: gb\_pr1:\*  
86: gb\_pr2:\*  
87: gb\_pr3:\*  
88: gb\_pr4:\*  
89: gb\_pr5:\*  
90: gb\_pr6:\*  
91: gb\_pr7:\*  
92: gb\_pr8:\*  
93: gb\_pr9:\*  
94: gb\_r01:\*  
95: gb\_r02:\*  
96: gb\_in4:\*  
97: gb\_pr10:\*  
98: em\_ba3:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

| Result No. | Score  | Query Match | Length | DB ID         | Description        |
|------------|--------|-------------|--------|---------------|--------------------|
| 1          | 4185.6 | 85.8        | 6582   | 88 AF093419   | AF093419 Homo sapi |
| 2          | 2816.2 | 57.7        | 7516   | 95 RNMUPP1    | AJ001320 Rattus no |
| 3          | 2718.2 | 55.7        | 6288   | 94 MMU131869  | AJ131869 Mus muscu |
| 4          | 1745.6 | 35.7        | 1768   | 93 HSMUPP1    | AJ001319 Homo sapi |
| 5          | 1433.4 | 29.4        | 1681   | 9 AX011714    | AX011714 Sequence  |
| 6          | 1213.2 | 24.9        | 2703   | 94 MMAR000168 | AF000168 Mus muscu |
| 7          | 614.4  | 12.6        | 169660 | 90 AL161449   | AL161449 Human DNA |
| 8          | 614.4  | 12.6        | 196607 | 79 AL353639   | AL353639 Homo sapi |

|    |       |      |        |    |           |             |
|----|-------|------|--------|----|-----------|-------------|
| 9  | 548.8 | 11.2 | 5033   | 85 | AB044807  | Homo sapi   |
| 10 | 409.4 | 8.4  | 3761   | 94 | AF060539  | Mus muscu   |
| 11 | 433.2 | 6.8  | 339    | 54 | G25381    | human STS E |
| 12 | 333.2 | 6.8  | 339    | 54 | G27450    | human STS S |
| 13 | 290.2 | 5.9  | 147739 | 63 | AC013514  | Homo sapi   |
| 14 | 285.2 | 5.8  | 149824 | 61 | AC009919  | Homo sapi   |
| 15 | 213.2 | 4.4  | 147739 | 63 | AC013514  | Homo sapi   |
| 16 | 208.4 | 4.3  | 4773   | 92 | HS2424747 | Homo sapi   |
| 17 | 187.2 | 3.8  | 5350   | 93 | HS2424747 | Homo sapi   |
| 18 | 185.6 | 3.8  | 4863   | 92 | HS2424748 | Homo sapi   |
| 19 | 178.4 | 3.7  | 149824 | 61 | AC009919  | Homo sapi   |
| 20 | 169.8 | 3.5  | 41381  | 5  | CB846119  | Caenorhab   |
| 21 | 168.4 | 3.5  | 42100  | 5  | CEC52A11  | Caenorhabd  |
| 22 | 120.6 | 2.5  | 3039   | 5  | AF274350  | Drosophila  |
| 23 | 120.6 | 2.5  | 3191   | 5  | AF132193  | Drosophila  |
| 24 | 120.6 | 2.5  | 3217   | 5  | AF103942  | Drosophila  |
| 25 | 120.6 | 2.5  | 30583  | 64 | AC017314  | Drosophila  |
| 26 | 120.6 | 2.5  | 257867 | 4  | AC005557  | Drosophila  |
| 27 | 120.6 | 2.5  | 303191 | 4  | AE003472  | Drosophila  |
| 28 | 105   | 2.2  | 143226 | 90 | AL353802  | Human DNA   |
| 29 | 105   | 2.2  | 143226 | 90 | AL353802  | Human DNA   |
| 30 | 99.4  | 2.0  | 128915 | 78 | AL356458  | Homo sapi   |
| 31 | 99.4  | 2.0  | 128915 | 78 | AL356458  | Homo sapi   |
| 32 | 88    | 1.8  | 61083  | 77 | AC090170  | Homo sapi   |
| 33 | 82    | 1.7  | 106    | 54 | G20352    | human STS A |
| 34 | 81    | 1.7  | 85173  | 90 | AL490787  | Human DNA   |
| 35 | 81    | 1.7  | 169329 | 81 | AL390787  | Homo sapi   |
| 36 | 78.8  | 1.6  | 162486 | 71 | AC044779  | Homo sapi   |
| 37 | 73.6  | 1.5  | 10763  | 89 | AF338650  | Homo sapi   |
| 38 | 70.2  | 1.4  | 4984   | 94 | AF175292  | Mus muscu   |
| 39 | 69    | 1.4  | 162486 | 71 | AC044779  | Homo sapi   |
| 40 | 67.4  | 1.4  | 3071   | 97 | HS032376  | Human chann |
| 41 | 67.2  | 1.4  | 8301   | 94 | AF169411  | Rattus no   |
| 42 | 66    | 1.4  | 1497   | 89 | AF233323  | Homo sapi   |
| 43 | 66    | 1.4  | 4424   | 93 | HS032376  | Homo sapi   |
| 44 | 66    | 1.4  | 7546   | 97 | HS032376  | Homo sapi   |
| 45 | 66    | 1.4  | 8043   | 93 | HS032376  | Homo sapi   |

## ALIGNMENTS

## RESULT 1

AF093419 6582 bp mRNA PRI 30-SEP-1998  
LOCUS Homo sapiens multi PDZ domain protein MUPP1 (MUPP1) mRNA, complete

ACCESSION AF093419  
VERSION AF093419.1 GI:3668409

SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 6582)  
AUTHORS Eng, L., Kravitsky, G. and Clapham, D.E.  
TITLE Human homolog of MUPP1 protein  
JOURNAL Unpublished

REFERENCE 2 (bases 1 to 6582)  
AUTHORS Eng, L., Kravitsky, G. and Clapham, D.E.  
TITLE Direct Submission  
JOURNAL Submitted (21-SEP-1998) Cardiology, Children's Hospital, 320  
Longwood Ave, Boston, MA 02115, USA

FEATURES  
source  
1. 6582  
Location/Qualifiers

/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/issue\_type="brain"  
/dev\_stage="fetus"  
1. 6582  
/gene="MUPP1"  
47..6175  
/gene="MUPP1"

BASE COUNT 2024 a 1331 c 1605 g 1622 t  
ORIGIN

Query Match 85.8%; Score 4185.6; DB 88; Length 6582;  
Best Local Similarity 93.6%; Pred. No. 0;  
Matches 4508; Conservative 0; Mismatches 14; Indels 294; Gaps 2;

212 aggttaaagtcataacttacttcttgagggaataaccacaagatgtgtgaatatcttaaaag 271  
1848 AAGTAATGCGATTAATCTTACTTGGGAAATATCCAAAGATGTGTGAATATCTTAAG 1907  
272 aactgcctataagaatgtaaatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 331  
1908 AACGCTATAGAGTACATAGT 1967  
332 cagaatgtgataagtcataacttcttgagggaataaccacaagatgtgtgaatatcttaaaag 391  
1968 CAGATTTGATAGCTGATGATTTATGATGATTTATGATGATTTATGATGATTTATGATGAT 2027  
392 tagtgatgtcattcagatgtcattcagatgtcattcagatgtcattcagatgtcattcagat 451  
2028 TAGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2087  
452 gtaagatgataagaatgtaaatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 511  
2088 GTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2147  
512 tagagctgag 571  
2148 TAGAGCTGAG 2207

/note="similar to Rattus norvegicus multi PDZ domain  
protein"  
/codon\_start=1  
/product="multi PDZ domain protein MUPP1"  
/protein\_id="AAC61870.1"  
/db\_xref="GI:3668410"



OY 572 ttgatccagcaagcacttgatataatcttcttgctgctggcgagcatctgtgaaa 631  
|||||  
Db 2208 TTGATCCAGCAAGCACTGTGATTAAACGTTCTTTGGTGGCTGGGCGATGCTGAAA 2267  
OY 632 agaatgagacactcttccctggtgagcagctcatgttctgaagcatgttcaacttgaaa 691  
|||||  
Db 2268 AGGATGGACGACTTCTTCTGGTGACCGACTCATGTTTGTAAACATGTTAACTTGGAAA 2327  
OY 632 aacagcagctctgaggaagcctgtagaagcctgaaaggagcacgctcaaggagctgtgaaa 751  
|||||  
Db 2328 ACAGCAGCTTGTGAGCAACTGTGAAGCACTGAAGGAGCAGCCGTCAAGGAGCTGTAGAA 2387  
OY 752 taagagtgcttaagccttacccttaccagaaagaggtatgttcttgtaaggagg 811  
|||||  
Db 2388 TAGGGGTGTAAAGCTTTACCCCTTTACCGAAGAAAGTTATGTTTCTGCTAAAGAGG 2447  
OY 812 attccttcttaaccacacacacccctgtgaaagcaggagctgtgacaaacccctct 871  
|||||  
Db 2448 ATTCTTCTCTACCCACCACTCTCTGTGAAGAAAGCAGGCGTGGCTGCAAAACCCCTCT 2507  
OY 872 tcaaggcctgacttggcctgtgtgagcaaaatgactgacttagtagatgaaatccacat 931  
|||||  
Db 2508 TCAGGCGTGACTTGGCTGTGGTGCAAAATGATGCTGACTTATGATGAAATCCACAT 2567  
OY 932 ttgagcttccatctctctctctgaaaatgacagatctactcttaactaaagcctctat 991  
|||||  
Db 2568 TTGACTTCCATCTCTCTCTGAAAATGACAGATCTACTCTAAGCTCTATTTAT 2627  
OY 992 ctctcctcagcagctctctctgtgtgagaaggcctgaactatgctctccctccatctc 1051  
|||||  
Db 2628 CTCTTCAATGAGGAGCTCTTGTGTGTATGGCTGAACTATGTTCTTCCATCATCTC 2687  
OY 1052 ctctcagagatgttatgtgaaaatctctgtgactcaagtaactgtatctgcatgtctcgg 1111  
|||||  
Db 2688 CTCTTAAGAGATGTTATTTGAAAATCTTGTGATCCAGTACTGATCTGATATGCTCTGG 2747  
OY 1112 aggaactctatcccaagatctcctcggaaagacagatgaaataaacctctgtgtgaaa 1171  
|||||  
Db 2748 AGGAACCTATATCCAGAACTCTCTGCAAGACAGAGATGAGATACCTTGTGGTGAAA 2807  
OY 1172 taagatgagggcctgtctctgacttaactataaagtaactacacacctgcaaatgtatgt 1231  
|||||  
Db 2808 TAAATATGGGCGCTCTCTGTGCTTACTATTAATGAGTACACACCTGCAAAATGCTATTTG 2867  
OY 1232 aacaaacatataatgtgaaacacacatagtgtagactgaaatctcaatccaagtgaag 1291  
|||||  
Db 2868 AACACAATATGAATGTGAAAACACAAATAGTGTGACTGAATCTCATTTACCAAGTGAAG 2927  
OY 1292 ttatatcaagtcgaagactctctctgtgtacccgagatccagctggaaggcctctgagc 1351  
|||||  
Db 2928 TTATATCAAGTGCAGAACTCTCTGTGTGTAACCGGATTCAGCTGGAAGGGCTCTGAGT 2987  
OY 1352 acctctgtgaacagagctccctcctgctgttaatgtcgaagtgtgtcaatgtctcaaaatgtat 1411  
|||||  
Db 2988 ACGTCTTGAACAGAGCTCCCTGGCTGTAAATGCTGAGTGTCTCATGCTTCAAAATGTAT 3047  
OY 1412 ctaaagaatctctgtgaaagactataataatagcaaaagcaatctcaagctagaatgaa 1471  
|||||  
Db 3048 CTAAAGAAATCTTTTGAAGGACTATTATATAGCAAAAGGCAATCTAGCCCTAGGAAATGA 3107  
OY 1472 caagttaagtcataaagaatgctgtgggagatgatacgttcgaagcaatcttaagtgagtg 1531  
|||||  
Db 3108 CAGTTAAGTGTATTAAGATGAGCTTGGGATGATGCTTGAAGCAATTATATCAATGAGAGTG 3167  
OY 1532 ccatatagtcagagatgagcgagatctgcatgtgagactgcatctgtgccatlaaagagtg 1591  
|||||  
Db 3168 CCATTAAGTCAGATGCGCGGATTTGCCATTTGGGAGCTGCAATCTTGTCCATTAATGAAGAT 3227  
OY 1592 ctacacatagtgtaaaccaatgcccagagcagagactatgtgtgaaagacatctctcatatg 1651  
|||||  
Db 3228 CTACCATGATGTTAAACCAATGCCAGGCGCAGAGCTATGTTGAAGAGACATCTCTCATTTG 3287

OY 1652 gccctgacataaaaaactactatgtgcctgcagaaacatttggaagagttcaaaatlaagct 1711  
|||||  
Db 3288 GCCCTGACATTAATAATTTACTTATGTGCTGCGAACAATTTGGAAAGATTCAAAATTAAGCT 3347  
OY 1712 tggagcaacaactctggaagatgaatgagcactgagatatttctctacacactgagagag 1771  
|||||  
Db 3348 TGGGACAACAATCTGGAAGAGTAATGGCACTGATATATTTTCTTCATTAACATGCGCAGAG 3407  
OY 1772 acatctcagaatctacagagcagagagagagaggtgtgaagaagacgagactcaaaaca 1831  
|||||  
Db 3408 ACATTCAGAAATTAACAGAGCGAGAAAGAGGAGAGGTGAAGAAAGCCGAACCTTCAAAACA 3467  
OY 1832 cagcatatagcaatggaatcagcccaaggcgggtggaactctcggagagagaaaccaacaat 1891  
|||||  
Db 3468 CAGCATATAGCAATTTGAATTCAGCCAGGCGGCTGGAATCTGGAGAGAACCAACCAAAAT 3527  
OY 1892 ccttagacatcagacatgttgtgtgagcagagagatgaggagctgcgtcaagcaatgagagag 1951  
|||||  
Db 3528 CCTTAGGCATCAGCATTTGTGTGTGACGAGGATGGGAGTGGGCTTAAGCAATGAGAGAG 3587  
OY 1952 tgaagaaggagcattctcatcaaaacatgtctcggaaagatagtcagctgtgcaaaatgaa 2011  
|||||  
Db 3588 TGATAGGGGGCATTTTTCATCAAAACATGTTCTGGAAGATAGTCCACCTGGCAAAATGGA 3647  
OY 2012 ccttgaacactggaagatgaatcgtgag ----- 2038  
|||||  
Db 3648 CTTGAACCTTGAGATGAAATGTAAGAGTGAAGTGAATGACCTCAGAGATGCAAGCC 3707  
OY 2039 ----- 2038  
Db 3708 ATGAACAAGCTGTGAAGCATTTGGAAGAGCAGCAACCTGTAGCTTATGTATACAGA 3767  
OY 2039 ----- 2038  
Db 3768 GCATTATAAACAGCAAGAAATCCCTTTCCTTCTGCTGCAACACTTTACCTTA 3827  
OY 2039 ----- 2038  
Db 3828 AGTACAACTTGACGACACTAACCCATTTGCTGACTCTCTACAAATCAAGCGCGACAMAG 3887  
OY 2042 caccagtcagtcagagtcagagtcagagagagagtcacatgtgtcagtgctgtcccccacc 2101  
|||||  
Db 3888 CACCAGTCAGTCAGAGTCAGAGCCAGAGAAAGCTTCATTTGTCAGTGTGCCCCACCCC 3947  
OY 2102 ctccctcagcctctgtccgaaatgagtgatgtatcacacacagctcatctgcaagcaaatct 2161  
|||||  
Db 3948 CTCTTTCAGCCCTTTCGCCGAATGGGTAGTGAACACACAGTCAATCTGCACCAAAATCT 4007  
OY 2162 cacaagaatgtgacaagaagagatgagttgtgtacagctggaagaaatatacagaagagct 2221  
|||||  
Db 4008 CACAAGATGTGACAAAGAGATGAGTTGTTACAGCTGGGAAAATATCAGAGAGCGTT 4067  
OY 2222 atggaacccctaaacagcgagctgtcatatgatatgaaatggaagagagtcagtagtgctg 2281  
|||||  
Db 4068 ATGGAACCTTAACAGCGGAGCTGCATATGATTTGAATGAGTGAAGGATCTAGTGTGG 4127  
OY 2282 gcttaagctctgtgtggaacaaagccgatacagaatgagtgctctcatagtgagagtg 2341  
|||||  
Db 4128 GCTTAAGTCTTCTGTGGAACAAAGACCGATCCAGATGAGTGTCTTCAATGAGGGGATTTG 4187  
OY 2342 atccaaatgagcctgcaagaaagaatgtgtcgaatgcaaatctcagaatgagctctcaga 2401  
|||||  
Db 4188 ATCCAAATGAGCTGCAGGAAAGATGTGCAATTCGAATTTCCAGATGAGCTTCTAGAGA 4247  
OY 2402 tcaatgtgcagatttatatggaagaagtcacagaaatgagcctcatcaatcaatlaaagtg 2461  
|||||  
Db 4248 TCAATGTGCAGATTTTATATGGAAGAGATCAATAATCCCTCATCAATTAATATGAG 4307  
OY 2462 cccctctcaagtgaaataatcttctacagaataaagaatgcaatgagtcagtagagcg 2521  
|||||  
Db 4308 CCCCTTCTTAAGTGAATAATTTTATCTAGAAATTAAGATGCAATGATGATGAGTGGCGG 4367  
OY 2522 taagtccgtgaagatgagtagaaccttgcctctcaactcagaataatctcaaaataag 2581

[illegible]

|    |      |  |      |
|----|------|--|------|
| Db | 5448 | TGGAAGTTGGAAAGATCTAAAG-----                                      | 5468 |
| OY | 3662 | gccaggtgagtgtaagcagccgtgtctcttcaacttccactctctgatacagtaacat       | 3721 |
| Db | 5469 | -----CTGGATCCAGTACAT   | 5483 |
| OY | 3722 | ctgagtcactctgaaagtagtctcaagaagaatgatctggcatctgaaatacaaggattaa    | 3781 |
| Db | 5484 | CTGAGTCACTGGAAAAGTAGCTCAAGAGAAATGCAATGGCATCTGAATTACAGGATTTAA     | 5543 |
| OY | 3782 | gaacagtcgaatgaaaaaaggccctactgactcactggyaaatcagatcgcgtgagag       | 3841 |
| Db | 5544 | GAACAGTCGAATATAAAAAAGGCCCTACAGTACATCGGGAAATCAGCATCCCTGGAGGAG     | 5603 |
| OY | 3842 | taaggcaacccactctggtgagtggtgcataattatgtcaatgtatgacccaactgnaagt    | 3901 |
| Db | 5604 | TAGGACGCCACTTGGTAGTGTGCTTAATTTATTTTCAAATGTCACCCCACTGGAGTTG       | 5663 |
| OY | 3902 | cagcacagaccccaaaaactcagaagtctggygatagatgtgtcacaactctgtgscatcca   | 3961 |
| Db | 5664 | CAGCACAGACCCCAAAAACCTCAGATGTTGGGATAGATGTGTACCACTGTGGCAATGCA      | 5723 |
| OY | 3962 | ctgggggcatgactcacaacccaagcgttaactactgtgaaaaatgcatctgtctccaatgt   | 4021 |
| Db | 5724 | CTGAGGCGATGACTCACACCCCAAGCAAGTAACTACTGTGAAAAATCATCTGGCTCCATGG    | 5783 |
| OY | 4022 | aaatgacagtgatgtgtctcgagggagacgtgagtggtgtcaagtgatacatcaagagccgt   | 4081 |
| Db | 5784 | AAATGCAAGTGTGTTGCTGGAGAGAGACGTAGTGTGTGTCACAGCTATACGACGAGCCTG     | 5843 |
| OY | 4082 | caagttccagttcttcttctcaactggtgcgtgacgtcaaccagtalattcaagaatgattag  | 4141 |
| Db | 5844 | CAAGTTCACAGTCTTTCTTTCTACTAGGGCTCAGTCAAGCAAGTATATTTACAGATCATTTTAA | 5903 |
| OY | 4142 | gacctctccaatgtaaatctattacactgagagagagaccagaatgctcttaggttcaagta   | 4201 |
| Db | 5904 | GACCTCCTCAATGTAAGTCTATTACCTAGAGCGAGAGCCAAATGGCTTAAGCTTCACTA      | 5963 |
| OY | 4202 | taatttgaggaatgtagcagagccctcatggygaacttaccatltatgtltaaaacagtgttg  | 4261 |
| Db | 5964 | TAGTTGGAGGATGTAGCGACGCCCTCATGSGAGACTTACCATTATGTATTAACAGAGTTTG    | 6023 |
| OY | 4262 | caaaaggagacagcctctcgaagaagagacgctctgaaaaaagggggaatcaagtatctgtcca | 4321 |
| Db | 6024 | CAAAAGGAGACGCTCTCGAAGAGAGGACGCTGTGAAAAAGGGCCATAGATCTTCTGTGCA     | 6083 |
| OY | 4322 | atggcagcagagctcgaagaaggtctaccactgaaagagcgtgttccatctcttaacagggcaa | 4381 |
| Db | 6084 | ATGGCGAGAGCTAAGAAAGAGTCAACCATTAAGAAAGAGCTGTGCCATCTTAAAGGACAA     | 6143 |
| OY | 4382 | aaggcactgtaacttgaatgtgttctctcttgaaatggtgtccagaaatgaaaccaaccaa    | 4441 |
| Db | 6144 | AAGCAGCTGTCACTTTGATGATGGTTCTCTCTTGAATTTGGCTGCAGAAATTAACCAACCAA   | 6203 |
| OY | 4442 | ccccctactaacctccactcagctaaagaagaatgcactggtgcctgpeaaatltttagtgcgt | 4501 |
| Db | 6204 | CCCCCTAGCTACCTCTCTACTGTAAAGAGATGACGTGGCTCGAGCAATTTTAAAGCTGTG     | 6263 |
| OY | 4502 | ttcagccgggtgtcttcaaaaactgttaggggggaaataacaacttaagttctcttcttccact | 4561 |
| Db | 6264 | TTTACGCGGGCTTTCAAAACTGTAAAGGGGGAAATTAACACTTAAGTTTCTTTTCTCATCT    | 6323 |
| OY | 4562 | agaaatgcttcttcttactgcaacaactacaatcatttctcttctctctgtcatlttgtga    | 4621 |
| Db | 6324 | AGAAATGCTTTCTTCTTACTGCACAACCTAAATCATATTTTCTTTCTTCTTGTGCAATTTTGA  | 6383 |
| OY | 4622 | acttaagagaaagaatatgtgtgtaggtgaaatccgttttatttattgtggygatatactaa   | 4681 |
| Db | 6384 | ACTTAAAGAAAGAAATATTGTGTAGTGGAATCTCGTTTTTATTGTGGAGATATCTAA        | 6443 |
| OY | 4682 | tgttttagtcacatggygcaagaattaltacatgtctaagctgtgttagtataaagaaga     | 4741 |
| Db | 6444 | TGTTTTTATGTACATAGGGCAAGAAATTTATTACGTCTAAGCGGTGTAGTATTAAAGAAGA    | 6503 |





|            |   |   |             |
|------------|---|---|-------------|
| QY         | 4142  | gaattcttcaatgtaagtcattcatcaactctagagagagaccacaatgcttaagcttcagta     | 4201        |
| Db         | 6077  | gccccccacacatgcttaaacacattaacctatgaccgagagacacatgagcttaaggcttcacaca | 6136        |
| QY         | 4202  | taattgagagatgtagagagccccaagagagacttaaccaattatgtttaaaacagtgctg       | 4261        |
| Db         | 6137  | tttgatgtaggctatagcagaccctcatatgagaccttaccattttatgtttaaacagtgctgc    | 6196        |
| QY         | 4262  | caaaaggagcaagcctctgaaagcagcagctctgaaaaagggcgatcaagatcatgctgtca      | 4321        |
| Db         | 6197  | caaaaggagcagcagccgacagagatggcgctgttaaaaggggggtatcagatcattgctgtca    | 6256        |
| QY         | 4322  | atggcagagagctctagaaaggggtccaccatgaagaagactgttgcaatcccttaacgagaca    | 4381        |
| Db         | 6257  | acgggcaaaagcttgaagggcggtacccatgacaaacactgttccatctttaagccgagaca      | 6316        |
| QY         | 4382  | aaggcagctcaactctgatgtctctctctgaaatgctgcgcagaaatgtaaccaaccaca        | 4441        |
| Db         | 6317  | agggcagctgacccctcatgctggttctctcttgaattggctgcagagccgaacac-agccag     | 6375        |
| QY         | 4442  | ccacctagctaacctccctactctgttaaagaagaatgcactgctgcctgcaaatlttatagctg   | 4501        |
| Db         | 6376  | ctagctggcccccacctcttactgtaaacggagaggaactgttc-----acatg              | 6420        |
| QY         | 4502  | ttcagccgggtctccaactctgaaagggggaataaacaacttaagtctcttcttcacatc        | 4561        |
| Db         | 6421  | acctgttattatggggagaaactacgcccggcccgagaa-----                        | 6456        |
| QY         | 4562  | agaaatgcttctccttactctgaacaactaacaatcttcttctctctctctgcatlttgta       | 4621        |
| Db         | 6457  | acacactgattgtgttctctaacacaaacagcagtgtttcttcttaccgtgcgcatgttcatag    | 6516        |
| QY         | 4622  | actc-----aaagaagaagaaatattgtgtgaagtgaatcctgcttcttcttggtgagat        | 4675        |
| Db         | 6517  | tcttatgcttcaaacagaaaggagaggtttgcagaggttaacctgatttatctc--tgaat       | 6575        |
| QY         | 4676  | atcatagtcttctgtagcacatggtgcagaatcatctacaatgcctgaagctgttagatataa     | 4735        |
| Db         | 6576  | atctnacaattttatnagcatgctgagac-agaattattgctatgccattttgtttagatnaga    | 6634        |
| QY         | 4736  | gaagaataatcttaagaactaaccaagaagaatggtcttcagtaag-ctagatgtaaaatg       | 4794        |
| Db         | 6635  | aaacaaaataaagcacaagtgtgacacagagagatggcttcagaaaaatttaagatataaagtg    | 6694        |
| QY         | 4795  | aaatataataataagaagaagaatctcctgggagttcaaaaaaaatgcccctcaattggcaa      | 4854        |
| Db         | 6695  | gaatttttgaaaaaaagaaagcagctct--tgaacttattatagaacttcccccaattctagaag   | 6751        |
| QY         | 4855  | ctctac 4859   |             |
| Db         | 6752  | ttctac 6756   |             |
| RESULT     | 3   |   |             |
| LOCUS      | MMU131869   |   |             |
| DEFINITION | MMU131869 6288 bp mRNA  | ROD   | 28-JUL-1999 |
| ACCESSION  | AJ31869   |   |             |
| VERSION    | AJ31869.1 GI:4150877  |   |             |
| KEYWORDS   | mpdz gene; multiple PDZ domain protein.                           |   |             |
| SOURCE     | house mouse.  |   |             |
| ORGANISM   | Mus musculus  |   |             |
| REFERENCE  | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; |   |             |
| AUTHORS    | Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus |   |             |
| TITLE      | Simpson,E.H., Sufoolk,R. and Jackson,I.J.                         |   |             |
| JOURNAL    | Identification, sequence, and mapping of the mouse multiple PDZ   |   |             |
| MEDLINE    | Genomics 59 (1), 102-104 (1999)                                   |   |             |
| REFERENCE  | 2 (bases 1 to 6288)   |   |             |
| AUTHORS    | Simpson,E.H.  |   |             |





Db 3057 CCCAGAGCTTTTGAAGGCCCCCTTACTATAGCAAAAGGCGCTCCAGCCTAGTATGA 3116  
Oy 1472 cagtagtgcataaagaatgagcttgaggatgacgcgtcgaagcatctcaatgagtg 1531  
Db 3117 CAGTAAGTCCCAATTAATGATGCGCTGCGGAGTGTGCGGAGCATTCATTCATGCGGCG 3176  
Oy 1532 ccatgaagcgaatgagcgaatgagcttgagctgagctgcatctgctcaatgaagag 1591  
Db 3177 CCATTAGTGTGATGCGCGGATGCTGTGGCGACTGCTTTTGTCCATTAATGAGAAAT 3236  
Oy 1592 ctaccatgaatgtaacaaatgcccagcagcagcagcagcagcagcagcagcagcagc 1651  
Db 3237 CCACCATCAGTCTACCAATGCCAGGCGCATGCTGAGAAAGACATTCATTAATG 3296  
Oy 1652 gccctgacaataaatactctatgctgctgcaagaacattggaagagtcgaataagc 1711  
Db 3297 GACCTGATATCAAAATTTACTTACGTGCTGCGAAGCATTTGGAAAGATTCAGAGTAAT 3356  
Oy 1712 tgggacaaacatctgaaagatgaatgacatgacatgatacttctcacaacagcagag 1771  
Db 3357 TTGGTCAACAGCTGGAAGATTAATGCACTGATATTTTCTTCATACACTGGCAGAG 3416  
Oy 1772 acatccgaatctaacagagcgaagagagagagagagagagagagagagagagagag 1831  
Db 3417 ACATTCCAGAACTCCAGAGCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3476  
Oy 1832 cagcaatagcaattggaatgagcagcagcagcagcagcagcagcagcagcagcagc 1891  
Db 3477 CTGCTTACAGAGCTGAGCGCCAGCCAGCGAGGTGGAATTTGGAGAGAACCAACAAAT 3536  
Oy 1892 ccttgagatcagcagctgtgtgagcagagagatgaggagctgagtaagcaatgagag 1951  
Db 3537 CTTGGGCAATCGCATGTTGTGTGCGAGAAATGGGAGCGCGCTGACCAATGGTGAGG 3596  
Oy 1952 tgaatagggagcattctcaacaacatgtctcgaagaatagcaactgcaaaatgaga 2011  
Db 3597 TGATAGGGGGCTTTTTCATCAAAACATGTTCTTGAAAGACAGCTGACCTGCAAAATGGA 3656  
Oy 2012 ccttgaacactggaatagaaatcgtag 2038  
Db 3657 CTCTGAGCCAGGAGACAGAAATTAATGAGGTGATGATGAGCCTCAGAGATGCAAGCC 3716  
Oy 2039 2038  
Db 3717 ATGAACAAGCTGTGAAGCCATTTCGAAAGACGCAACCTGTAGTGTATGTACAGA 3776  
Oy 2039 2038  
Db 3777 GCATTATAACAGACCAAGAAATCCCTTTCCTTCCTTGGCGCACAGCCTTTACCC 3836  
Oy 2039 2041  
Db 3837 AGTACAGCTTACGACACTAACCCATTTCAGACTCTCTCCAGCTCAGCACTGACAG 3896  
Oy 2042 caccacagcagcagagatcagagcagagagagagagagagagagagagagagagag 2101  
Db 3897 CACCCAGCCAGCTGAGAAATCTGAGACAGAAAGCCTGCAATGTGCAATGTTCTCCCTCT 3956  
Oy 2102 ctccctcagccttgcgaatgagcagcagcagcagcagcagcagcagcagcagcagc 2161  
Db 3957 CTCCTTCAAGTGTTCAGAAATGGGATGATGATGACAGCAGCAGCTGCAAGCGAGTCT 4016  
Oy 2162 cacaagaatgagcaaaag 2221  
Db 4017 CAGAGAGTGAAGCAAAAGAGATGAGTGTGTTACAGCTGGAAGAAATATCCAGAGCGCT 4076  
Oy 2222 atggaacacttaacagagcagcagcagcagcagcagcagcagcagcagcagcagcag 2281  
Db 4077 ATGGAGCCTGACAGGTGACGCTCCATGTGATGAGCTGAGAAAGGCGAGACGGGTTGG 4136  
Oy 2282 gctaagctctgagcaaaagagagagagagagagagagagagagagagagagagag 2341

Db 4137 GTCTAGTCTTGCTGGAAACAAAGACCCAGCAGATGATGTATTAGTGGGATTCG 4196  
Oy 2342 atccaaatgagctgcaagaaagaatgctgcaatgcaatgcaatgcaatgcaatgca 2401  
Db 4197 ATCCCACTGAGACAGCGGAGAGATGGCCGACTACATTCGACAGAGCTTTAGAGA 4256  
Oy 2402 tcaatgcaatgcaatgcaatgcaatgcaatgcaatgcaatgcaatgcaatgcaatgca 2461  
Db 4257 TCAATGGCCAAATATTGATGAGCAGAGATGATGATGATGATGATGATGATGATGAT 4316  
Oy 2462 cccctcctaagatgaataatcttcaagaataaagaatgcaatgcaatgcaatgcaatg 2521  
Db 4317 CTCCATCTTAATAAATAAATTTTATCAGAAATGACAGATGACATGAAAGAGAGCGCG 4376  
Oy 2522 tatgtccctggaatgcaatgcaatgcaatgcaatgcaatgcaatgcaatgcaatgca 2581  
Db 4377 TATGTCTGGGATTCAGACAGCTCTCCGCTTCACCTCAGACAGCTCTCAGAACAGG 4436  
Oy 2582 agacagagcacaatgtaactctgagcagcagcagcagcagcagcagcagcagcagc 2641  
Db 4437 AGGTGGAACATGACAGTACTGATGATGATGATGATGATGATGATGATGATGATGAT 4496  
Oy 2642 tgaacatctgagcttcccaagagatcagggggttgggtatgctacagcgaagaag 2701  
Db 4497 TGTACACGTGAGCTTCTTAAGATCAAGAGAGTTAGGCTATGCTATCTGTGAGGAAG 4556  
Oy 2702 ataacatgagtgatcatcaataagagcttaacagagcagcagcagcagcagcagc 2761  
Db 4557 ACACATCAACGAGATCATGATGAGAGCTTAACGAGCAGCGGAGAGCAGCAAGAGAG 4616  
Oy 2762 gacagcagaatgagagatcagcagcagcagcagcagcagcagcagcagcagcag 2821  
Db 4617 GAGGCTCAACCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4676  
Oy 2822 ctatgaaagatgtaatgagccttgaagaacagcaagaatgcaatgcaatgcaatgca 2881  
Db 4677 CTGTTGAAGAATTCATGAGCTTGAAGAGAGCAAGCAAGCAAGCAAGCAAGCAAGCA 4736  
Oy 2882 atgtgagagatcagatcccaagcagcagcagcagcagcagcagcagcagcagcag 2941  
Db 4737 GAGCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4796  
Oy 2942 aagaagacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 3001  
Db 4797 GGAAGACACTCCAGACTCTGCAAGTCCA-----CTCCAGACTGGAACCCA 4847  
Oy 3002 tccgaatacagaagcagatcatcaacacagcaatcttgcctcagcagcagcagcagc 3061  
Db 4848 TCCCAAGTCCAGAGAGTCTCCACACAGCAGAGTGTGCTGACCTGACCTGACCTGCC 4907  
Oy 3062 ccatatccctgctgagcaaaacacacacacacacacacacacacacacacacacacac 3121  
Db 4908 CCATATCCCAAGCTGTGAGCAACAATGAGATTTCCAAAGCCAAAGAGCCTGGGAC 4967  
Oy 3122 tgaagcagctgagagagcagcagcagcagcagcagcagcagcagcagcagcagcag 3181  
Db 4968 TGAATATTGTTGGGCTGACACACAGCTCTGGGTCTATTATTATCCATGAGTTTAA 5027  
Oy 3182 aagaagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 3241  
Db 5028 AAGAGGAGACAGCGTGAAGATGAGACTATGGCTGAGACCAAGATTTAAGAGTAA 5087  
Oy 3242 atggaatgacttgaggaagcagcagcagcagcagcagcagcagcagcagcagcagc 3301  
Db 5088 ATGGATGACTTGGGAAAGGCTTACACATGATGAAGCAATATGCTCCAGAGAGAGCC 5147  
Oy 3302 cagagagctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 3361  
Db 5148 CTCGAAGAGTACGCTGAGCTCTACCGAGATGAGCCCATCAAGAGAGAGAGTGTCT 5207  
Oy 3362 gtgaacactcactatgagcagcagcagcagcagcagcagcagcagcagcagcagc 3421  
Db 5208 GTGATACCTTACATGAGTGCAGAAAGAGCCAGGCAAGAGCTTGGTGTGATGATG 5267





```

Oy 3412 ttaagatctgttgcgaagaacagatctggaagatctgttgcgaacatctgcaaga 3471
Db 361 ttaagtattgttgtaaaagaaacatctgagatctgttgcgaacatctgcaaga 420
Oy 3472 ggaattgcagatcccgatggaagactgataccagagacagatatattgttgaatgg 3531
Db 421 ggaattgcagatcccgatggaagactgataccagagacagatatattgttgaatgg 480
Oy 3532 gaagacgttctgaatgcttcccaagaagcggttgcgcttctgtcaatgttccctagg 3591
Db 481 gaaagcgttctgaatgcttcccaagaagcggttgcgcttctgtcaatgttccctagg 540
Oy 3592 aagaacacttgcgaagctggaagaatcaaaagctgttccatctcattcagaagaaggcca 3651
Db 541 aagaacacttgcgaagctggaagaatcaaaagctgttccatctcattcagaagaaggcca 600
Oy 3652 tctcaaacacccagctgagtggaagcgagcctgtcttcttccatttccatctctgga 3711
Db 601 tctcaaacacccagctgagtggaagcgagcctgtcttcttccatttccatctctgga 660
Oy 3712 tccagtaacatctgagatcactggaagaatagctcaagaagaatgcatctgcaata 3771
Db 661 tccagtaacatctgagatcactggaagaatagctcaagaagaatgcatctgcaata 720
Oy 3772 cagggattaagaacagctggaagaagaagcgccctactgactcattggaatcagcatc 3831
Db 721 cagggattaagaacagctggaagaagaagcgccctactgactcattggaatcagcatc 780
Oy 3832 gctggaagagtagtgcagccacttgtgtgtcctataattatgtcaatgtagcaca 3891
Db 781 gctggaagagtagtgcagccacttgtgtgtcctataattatgtcaatgtagcaca 840
Oy 3892 actggaagctgcagcagacccaactcagaatttgggagataagatgtcaccatctgt 3951
Db 841 actggaagctgcagcagacccaactcagaatttgggagataagatgtcaccatctgt 900
Oy 3952 ggcacatccactgagagctgacacaccccaagcagtttaaccttcaaaatcatct 4011
Db 901 ggcacatccactgagagctgacacaccccaagcagtttaaccttcaaaatcatct 960
Oy 4012 ggtccatctgaaatgcaagtggtgtcgtggaagagacgttgaatgtgtcagagtcacat 4071
Db 961 ggtccatctgaaatgcaagtggtgtcgtggaagagacgttgaatgtgtcagagtcacat 1020
Oy 4072 cagggactgcgaagcttcttcttccatctgagcgtcagacccaagatcttcaag 4131
Db 1021 cagggactgcgaagcttcttcttccatctgagcgtcagacccaagatcttcaag 1080
Oy 4132 gatgatttagagccctcccaatgtaagctatatacactagaagcagaagacagatgctta 4191
Db 1081 gatgatttagagccctcccaatgtaagctatatacactagaagcagaagacagatgctta 1140
Oy 4192 ggtctcagatagcttgcagagatagcagccctcactggaagacttaaccaattatgtaaa 4251
Db 1141 ggtctcagatagcttgcagagatagcagccctcactggaagacttaaccaattatgtaaa 1200
Oy 4252 acaagtgctgcaaaagcgagcctcctggaagcgaagcgtctgaagaagcgatcagatc 4311
Db 1201 acaagtgctgcaaaagcgagcctcctggaagcgaagcgtctgaagaagcgatcagatc 1260
Oy 4312 attgtctcaatgagcagatctagaagaagctacccatgaagaagcgtctgcaatcct 4371
Db 1261 attgtctcaatgagcagatctagaagaagctacccatgaagaagcgtctgcaatcct 1320
Oy 4372 aaacggacaagaagcagctgcaatctgtgtcttcttgaatctgtgctgcagaatgta 4431
Db 1321 aaacggacaagaagcagctgcaatctgtgtcttcttgaatctgtgctgcagaatgta 1380
Oy 4432 acaaacccaacccctagctcactcctactctgaagaagaatgcatgctcctgaatatt 4491
Db 1381 acaaacccaacccctagctcactcctactctgaagaagaatgcatgctcctgaatatt 1440
Oy 4492 ttatctgttgcagcgggtcttcaaaactgtaagggggaataaacttaagtttctt 4551

```

```

Db 1441 ttaactgtgtcacccggcttcaaaactgtaggggggaataaacttaagtttctt 1500
Oy 4552 ttctcactagaaatgcttccctactcagacaactcaacatcatttcttctcttg 4611
Db 1501 ttctcactagaaatgcttccctactcagacaactcaacatcatttcttctcttg 1560
Oy 4612 catcttggaacttaagaagaagatatctgtgaagtgaatcttcgtttatttttg 4671
Db 1561 catcttggaacttaagaagaagatatctgtgaagtgaatcttcgtttatttttg 1620
Oy 4672 agatatcctaatcttctgtagtaccatggaagaaattatcaatgtaagctgttagta 4731
Db 1621 agatatcctaatcttctgtagtaccatggaagaaattatcaatgtaagctgttagta 1680
Oy 4732 taaagaagaataatctcctaagcttaacccaagaagaatgctcagtaagttagatgaata 4791
Db 1681 taaagaagaataatctcctaagcttaacccaagaagaatgctcagtaagttagatgaata 1740
Oy 4792 atgaaatataataaagaagaataa 4817
Db 1741 atgaaatataataaagaagaataa 1766

RESULT 5
AX011714 1681 bp DNA PAT 06-SEP-2000
LOCUS Sequence 112 from Patent WO955858.
DEFINITION AX011714
ACCESSION AX011714
VERSION AX011714.1 GI:998238
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 1681)
AUTHORS Schmitt,A., Specht,T., Dahl,E., Hinzmann,B., Rosenthal,A. and
Pilarsky,C.
TITLE Human nucleic acid sequences obtained from pancreas tumor tissue
JOURNAL Patent: WO 955858-A-112 04-NOV-1995;
BERNDT ARMIN (DE); SPECHT THOMAS (DE); DAHL EDGAR (DE); HINZMANN
BERND (DE); ROSENTHAL ANDRE (DE); METZGEN GES FUER GENOMFORSCHUN
(DE); PILARSKY CHRISTIAN (DE)
FEATURES
source location/Qualifiers
1..1681
/organism="Homo sapiens"
BASE COUNT 498 a 330 c 445 g 408 t
ORIGIN
Query Match 29.4%; Score 1433.4; DB 9; Length 1681;
Best Local Similarity 99.1%; Pred. No. 0;
Matches 1462; Conservative 0; Mismatches 11; Indels 2; Gaps 2;
Oy 2106 ttcaagccttgcgaagaatgagtgatgacacagaagcatctggaagaataatccaca 2165
Db 1 ttcaagccttgcgaagaatgagtgatgacacagaagcatctggaagaataatccaca 60
Oy 2166 agatgtgcaaaagagatgagttgttaccagcctggaaaaaatcacagaagcgtttag 2225
Db 61 agatgtgcaaaagagatgagttgttaccagcctggaaaaaatcacagaagcgtttag 120
Oy 2226 aaccttaacagcgagctgcatatgatactggaagcgaagaagctcagtggttggcct 2285
Db 121 aaccttaacagcgagctgcatatgatactggaagcgaagaagctcagtggttggcct 180
Oy 2286 aagcttgcgggaacaaagacagatcagaatgagtggtctcagaagggagatgattcc 2345
Db 181 aagcttgcgggaacaaagacagatcagaatgagtggtctcagaagggagatgattcc 240
Oy 2346 aatgagcctgaggaagaagatgctgcatctgcaatcagatgagctcctagatcaa 2405

```



Db 121 AACTCCCAACACCTCCGACATCCCA-----GCTCCAGACCTGGAAACCATCCCAACT 171  
 Qy 3010 acaagacagatcacaacacagcaatlttgccttcgacacccgacacccatc 3069  
 Db 172 ACAAGAGCTCCCTCCACACACAGCTCTTGGTTCTGACCTGCCATGCCCATCATC 231  
 Qy 3070 cctgctcgtgaaacacacccatcgaatlttccaaaggcgagacggcgctgagcacc 3129  
 Db 232 CCAGCGCTGGAACAACAATTTGAGTTTCCAAAGGCCAAAGCGCTGGAGCTGATTT 291  
 Qy 3130 gttgggggtcagacacacgctgctggtgcttattatccagaagttaagaaga 3189  
 Db 292 gttgggggtcagacacacacgctgctggtgcttattatccagaagttaagaaga 351  
 Qy 3190 gcaagcgtgaaagatggaagactcgtgctgagatcagaacttgaaggatgaat 3249  
 Db 352 GCAGCGCTGGAAGATGGAAGACTATGGCTGAGACCAATTTTAAAGCTAAATGCGATT 411  
 Qy 3250 gaactgaggaaggccacacatgatatgagaatcaatgctcctgagagacgcaagaga 3309  
 Db 412 GACTTGGCAAGAGCTACACATGATGAAGCAATCAATGCTTGAAGGACGCTCAAGA 471  
 Qy 3310 gtcgcctcagacactcagagatgagggcccaatacaagagagagagatgtgagacac 3369  
 Db 472 GATCGGCTGACGCTCTACAGAGATGAGGCCCATACAAAGAGAGAGATGCTGTGATAC 531  
 Qy 3370 ctaactatg-----agctcgaagaagacgggaaaggcctagatlaaglatgtt 3423  
 Db 532 TTCCACCATGAGCTGAGCTGAGAGAGAGCGGCAAGAGCGCTGGTGGTATGTTT 591  
 Qy 3424 ggttaaaagaagatcagatgaglatgtgtgaagatgtcaaaaggagaaatgagat 3483  
 Db 592 GCGAANAAGAAATGACACTGTGAGTATTGTATACAGACTTGTCAAGGAGGACATGACAG 651  
 Qy 3484 ccgagatgagaagatcagatccagagagacagatataatgtgagatgagagacgttc 3543  
 Db 652 GCCGATGGAGACTGATGCAAGGGGACAGATTTTAAAGTGAATGAGAAAGATGTCCT 711  
 Qy 3544 aatgcctcccaagagagagatgctgctgctcgaagatgttcctcgaagacagtaact 3603  
 Db 712 CATGCCACACAGAGAGACTGCTGCTGCCCTGCTAAAGTTCCTTGAAGTAAACCTG 771  
 Qy 3604 gaagttagaagaacacagctgcatccatccatccagagagagagacatcccaaccac 3663  
 Db 772 GAGGTTGGAAAGTCAAACTGCTCCCATTCACACTCAGAACGGAGGCTTCTCAAAACAG 831  
 Qy 3664 caggtgagtgaaagcagcctgctccttcacatlttccactcctcgtgacagatcatc 3723  
 Db 832 CAGGTGAGTGAAGAGAGCTGCTGCTCATCTTCACTCCCTCCACTTCTGGAATTAATCATCA 891  
 Qy 3724 gagtcacgtgaaagtgcacaaagaatgcatgtgcacgtgaatcagagatgaaga 3783  
 Db 892 GAGTCTTTGGAAGTAACTCGAAGAAATGCAATTCGAAATTCGAAATTCGAAATTCG 951  
 Qy 3784 acagtcgaagtgaagaaggccctactgactcaactcagagatcagacgtcgtgagagaga 3843  
 Db 952 ACAGTGAATTAATAAAGGGGCTGCTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1011  
 Qy 3844 ggcagcccaactgtgtgactgctcctatattatgtcagatgacacacacacacacacac 3903  
 Db 1012 GCGAGCGCTCGAGATGCTCCCTATATTATTTGATGATGATGATGATGATGATGATGAT 1071  
 Qy 3904 gcaagac 3963  
 Db 1072 GCTCAAAACCAAAACCAAAACCAAAACCAAAACCAAAACCAAAACCAAAACCAAAACCA 1131  
 Qy 3964 ggggggagatcacaac 4023  
 Db 1132 CATGGATACATCAC 1191  
 Qy 4024 atgcagagtggtgtgctgagagagacgtgagatgtgtacacagatcatcatcagagacac 4083

Db 1192 GTACAGAGTGTTGCTGGAGAGATGAGTGTGTGTCAGCGGTATCAGCAAGAACTTGCC 1251  
 Qy 4084 agttccagtccttcttcacgtggtcagcgtcacaacagatatttcaagatgagga 4143  
 Db 1252 AATCTTGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1311  
 Qy 4144 ccttcctcaatgtaagtctcttcaactagagcagagacagatgagcttgaagttata 4203  
 Db 1312 CTTCCACAGTCTTAAGCAATTCAGTTCAGACGACGACGATGCTTACGCTTACGATTT 1371  
 Qy 4204 gttgagagatagcagccctcagagatcagacacacacacacacacacacacacacacac 4263  
 Db 1372 GTAGGGGCTATGCGAGCCCTCATGAGACTTACCAATTAATGTAAGAAACAGTTTGA 1431  
 Qy 4264 aaggagacagcctcagagacagagcgtcgaagaaggcgatcagatcagatcgtcgaat 4323  
 Db 1432 AAGGGAGCAGCAGCAAAATGAGCTTAAAGAGGGGAGATCAGATTCCTGCTGCAAT 1491  
 Qy 4324 gggcagagctagaaagagatcaccacagagagacgtcgtcagatccttcaagacacaa 4383  
 Db 1492 GGGCAAGTCTTAAGAGAGTGAACCATGAGAGAGCTGTTGCCATCTTCAAGAGACAAAG 1551  
 Qy 4384 ggcactgcatctgagatgtctcctcctcctcctcctcctcctcctcctcctcctcctc 4443  
 Db 1552 GGCACCGTCAACCTCATGTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1611  
 Qy 4444 cctagtcacacccctcactcgttgaagaagatgacatgctgctcagacatlttgcgtgt 4503  
 Db 1612 ACTGGCTCCCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1661  
 Qy 4504 cagcgggtcctcaaaacgtgagggagagatcaacacacacacacacacacacacacac 4563  
 Db 1662 GGTCCGAGAGGCTACAGTAAAGGAGCGAGA-----AAAAAAAAAAAAAAAAAAAA 1694  
 Qy 4564 aaatgcttccctcactcagac 4618  
 Db 1695 CAGAGTGTGTTACTACAGCAGCAACATCATTTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1754  
 Qy 4619 tgaacttaagaagaagatattgtgtgagtgagatcctgcttcttcttcttcttcttct 4678  
 Db 1755 TTACTCAAAAGAAAGATATTTCATGATTAACCTTCAATTTATCT-CCAGCAATTAAC 1813  
 Qy 4679 taatgttctgacacagggcagaagaatatacagtcgaagctggtcgttgaagaaga 4738  
 Db 1814 TAAACATTTTAAGTCAACGCGAGCAA-AAATTTATATGTTCACTTGTATGAGAGAAAC 1872  
 Qy 4739 agataatcctaagctaaccaa 4760  
 Db 1873 AAAATGATACAAAGTTAGGCAA 1894  
 RESULT 7  
 AL161449/c 169660 bp DNA 09-JUL-2000  
 LOCUS Human DNA sequence from clone RP11-382H24 on chromosome 9p22.1-23,  
 DEFINITION complete sequence.  
 ACCESSION AL161449  
 VERSION AL161449.7 GI:8894259  
 KEYWORDS HTG.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 169660)  
 AUTHORS Sehra,H.  
 TITLE Direct Submission  
 JOURNAL Submitted (09-JUL-2000) Sanger Centre, Hinxton, Cambridgeshire,  
 CB10 1SA, UK. E-mail enquiries: humbrey@sanger.ac.uk  
 COMMENT On Jul 1, 2000 this sequence version replaced gi:8653807.  
 During sequence assembly data is compared from overlapping clones.  
 Where differences are found these are annotated as variations  
 together with a note of the overlapping clone name. Note that the

variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence has been finished according to sequence map criteria as follows. An attempt is made to resolve all sequencing problems, such as compressions and repeats, but not necessarily within known annotated human repeat sequence elements (e.g. Alu). Where the sequence is ambiguous, there is an annotation using the 'unsure' feature key.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases:

Em, EMBL; Sw, SWISSPROT; Tr, TrEMBL; Wp, WormPEP; Information from the WormPEP database can be found at <http://www.sanger.ac.uk/projects/C-elegans/wormpep> This sequence was generated from part of bacterial clone contigs of human chromosome 9, constructed by the Sanger Centre Chromosome 9 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr9>

RP11-382H24 is from the library RP11-1.2 constructed at the Roswell Park Cancer Institute by the group of Pieter de Jong. For further details see <http://bacpac.med.buffalo.edu/>

VECTOR: pBAC3.6

This sequence is the entire insert of clone RP11-382H24. The true left end of clone RP11-664D14 is at 135950 in this sequence. The true right end of clone RP11-187K14 is at 53851 in this sequence.

Location/Qualifiers

1. 169660

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/chromosome="9"

/map="P22.1-23"

/clone="RP11-382H24"

/clone\_11b="RP11-1.2"

1659..1724

/note="12 copies 28 mer 92% conserved"

2068..2401

/note="LIMC4 repeat: matches 7453..7834 of consensus"

2409..2828

/note="15 copies 28 mer 62% conserved"

2429..2848

/note="7 copies 60 mer 65% conserved"

2452..2801

/note="175 copies 2 mer 61% conserved"

4621..4865

/note="MIR repeat: matches 1..261 of consensus"

5200..5294

/note="L2 repeat: matches 2578..2670 of consensus"

6031..6654

/note="LIPAL2 repeat: matches 5583..6152 of consensus"

7431..7737

/note="AluSc repeat: matches 1..307 of consensus"

7763..8103

/note="MLT1A1 repeat: matches 205..536 of consensus"

8379..8422

/note="22 copies 2 mer 86% conserved"

9774..9835

/note="LIM47 repeat: matches 6223..6284 of consensus"

10099..10533

/note="MLT1C repeat: matches 9..466 of consensus"

10789..11105

/note="AluSc repeat: matches 1..309 of consensus"

11109..11240

/note="MER5A repeat: matches 62..189 of consensus"

11224..11267

/note="MER5B repeat: matches 138..178 of consensus"

14993..15030

/note="19 copies 2 mer 97% conserved"

16682..16772

/note="MLT2D repeat: matches 1..90 of consensus"

16759..16999

/note="MLT1A1 repeat: matches 1..241 of consensus"

17339..17638

/note="AluY repeat: matches 4..304 of consensus"

repeat\_region

17677..18235

/note="L2 repeat: matches 2009..2750 of consensus"

repeat\_region

18882..18943

/note="MLT1J repeat: matches 411..467 of consensus"

repeat\_region

18986..19291

/note="AluSc repeat: matches 1..307 of consensus"

repeat\_region

19450..19635

/note="MIR repeat: matches 59..248 of consensus"

repeat\_region

20230..20295

/note="13 copies 2 mer at 75% conserved"

repeat\_region

21640..22131

/note="MLT1H repeat: matches 40..547 of consensus"

repeat\_region

22141..22256

/note="MLT1H repeat: matches 236..352 of consensus"

repeat\_region

22708..22940

/note="LIP3A3 repeat: matches 5401..5632 of consensus"

repeat\_region

22941..23436

/note="LIP3A3 repeat: matches 5641..6146 of consensus"

repeat\_region

26193..26290

/note="MIR repeat: matches 97..193 of consensus"

repeat\_region

27837..27899

/note="L2 repeat: matches 2641..2708 of consensus"

repeat\_region

28653..28739

/note="ALUJO/FRAM repeat: matches 202..289 of consensus"

repeat\_region

29135..29237

/note="MIR repeat: matches 97..194 of consensus"

repeat\_region

29262..29788

/note="L2 repeat: matches 2198..2742 of consensus"

repeat\_region

29958..30391

/note="MLT2B repeat: matches 1..444 of consensus"

repeat\_region

30392..30590

/note="HERV1 repeat: matches 5456..5654 of consensus"

repeat\_region

30591..30631

/note="MLT2B repeat: matches 1..41 of consensus"

repeat\_region

30633..31387

/note="L2 repeat: matches 1156..1991 of consensus"

repeat\_region

31401..31876

/note="MLT1H repeat: matches 13..516 of consensus"

repeat\_region

31936..32142

/note="L2 repeat: matches 929..1149 of consensus"

repeat\_region

32790..33087

/note="AluSc repeat: matches 1..297 of consensus"

repeat\_region

35089..35209

/note="LIM4 repeat: matches 4928..5049 of consensus"

repeat\_region

35239..36006

/note="LIM1 repeat: matches 5352..6103 of consensus"

repeat\_region

37784..37985

/note="LIMC4 repeat: matches 5695..5897 of consensus"

repeat\_region

38615..38841

/note="MIR repeat: matches 3..241 of consensus"

repeat\_region

38875..39054

/note="13 copies 60 mer 76% conserved"

repeat\_region

38900..39039

/note="5 copies 28 mer 95% conserved"

repeat\_region

39052..39352

/note="AluSc repeat: matches 1..302 of consensus"

repeat\_region

39385..39681

/note="AluSc repeat: matches 1..298 of consensus"

repeat\_region

40167..40470

/note="AluSc repeat: matches 5..309 of consensus"

repeat\_region

40879..41165

/note="AluDb repeat: matches 1..286 of consensus"

repeat\_region

44053..44310

/note="AluY repeat: matches 53..308 of consensus"

repeat\_region

44543..45059

/note="MLT1H repeat: matches 64..535 of consensus"

repeat\_region

45124..45229

/note="L2 repeat: matches 2160..2272 of consensus"

repeat\_region

47382..47705

/note="ALUJO repeat: matches 1..312 of consensus"

repeat\_region

48964..49265

/note="AluSc repeat: matches 11..312 of consensus"

repeat\_region

50925..51071

```

repeat_region /note="L2 repeat: matches 2579. .2750 of consensus"
51236. 51431
/note="WERB repeat: matches 9. .210 of consensus"
51507. 51613
repeat_region /note="MIR repeat: matches 53. .146 of consensus"
51614. 52055
repeat_region /note="MSTR repeat: matches 5. .426 of consensus"
52056. 52103
/note="MIR repeat: matches 3. .53 of consensus"
52372. 53211
repeat_region /note="L1MD2 repeat: matches 4972. .5856 of consensus"
53204. 53521
repeat_region /note="L1MD2 repeat: matches 4649. .4975 of consensus"
53617. 53744
/note="64 copies 2 mer aa 58% conserved"
54114. 54573
repeat_region /note="MLTIC repeat: matches 1. .466 of consensus"
55083. 55144
repeat_region /note="31 copies 2 mer ca 71% conserved"
55981. 56112
repeat_region /note="MLTII repeat: matches 3. .136 of consensus"
57246. 57712
repeat_region /note="MLTID repeat: matches 1. .502 of consensus"
58718. 59833
repeat_region /note="L1MB1 repeat: matches 5049. .6170 of consensus"
60653. 61120
/note="L1MA9 repeat: matches 5687. .6139 of consensus"
62690. 62872
repeat_region /note="L1MA9 repeat: matches 6130. .6308 of consensus"
64488. 64765
repeat_region /note="AlusX repeat: matches 1. .286 of consensus"
65454. 65631
repeat_region /note="L1 repeat: matches 4470. .4650 of consensus"
65633. 66246

```

```

Query Match 12.6%; Score 614.4; DB 90; Length 169660;
Best Local Similarity 99.8%; Pred. No. 1e-139;
Matches 615; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

OY 4265 agggagcagcctctgaagcagcgtctgaagaaggagcagatcagatcgtctgaatg 4324
|||||
DB 146842 AGGAGAGCAGCCCTCTGAMAGCGACGCTGMAAAGGGCGCATGAGATTCATTCGCAATG 146783
|||||
OY 4325 ggcagagctagaagagagcaccacatgaagagcctctgacatccttaagagcagaag 4384
|||||
DB 146782 GGCAGAGCTTAAAGAGAGTCAACCATGAGAGAGCTGTTCCTTTAAACGACAAAG 146723
|||||
OY 4385 gcactgacacttgatgctctctctgaattgctgccaagaattgaacccaacc 4444
|||||
DB 146722 GCACGTGCTACTTGAATGCTTCTCTGAAATGGCTGCCAGAAATGAACCAACCAACC 146663
|||||
OY 4445 ctgactcaccctccactcgtaaagagaaatgcactggtcctgacaatttttaagctgtgc 4504
|||||
DB 146662 CTAGCTTCACCTCTACTGTAAGAGAAATGCACTGGCTCGACAAATTTTATCTGTGTC 146603
|||||
OY 4505 agccgggtcttaaaactgtaagggggaataatacctaagttcttcttcacatctaga 4564
|||||
DB 146602 ACCCGGGCTTCAAAACGTAGGGGGGAATAACACTTAAGTTCTTTTTCATCTAGA 146543
|||||
OY 4565 aatgctcttcttctgacaacacttaacatcatcttcttctctctgcaatttgtaagt 4624
|||||
DB 146542 AATGCTTCTTCTACTGACACCACTAACATCATTTTCTTTCTTCTTGCAATTTTGTAAGT 146483
|||||
OY 4625 taagaagaagaatacttgtaagtgtaactcgttttatttatttgtagagatacttaagt 4684
|||||
DB 146482 TAAAGGAAGAATATTGTGTAGGTGATCTGTTTATTTATTTGTGAGATATCTAATCT 146423
|||||
OY 4685 ttgtgagtcacatgggcaagaattattacatgctaaagctggttagtataaagaagata 4744
|||||
DB 146422 TTGTGAGTCACATGGCAAGAAATTTATCATCTAAGCTGCTTGTAGATTAAGAAACATTA 146363
|||||
OY 4745 ttctaaagtaacaaagaagaatgctcagtaagcttaagtaagaataataataa 4804
|||||

```

```

DB 146362 TTCTAAGCTAACCAAGAAATGGCTTCAGTAATTAAGATGAATAATGAATATATAA 146303
OY 4805 ataagaagaagaatctcggggaglttaaaaaaatgctcaatttggaattctctct 4864
|||||
DB 146302 ATAAAGAAAGAAATCTCGGGGAGTTTAAAAAAATGCTCATTTGGCAATCTACTCTCT 146243
OY 4865 ctccccaccccaact 4880
|||||
DB 146242 CTCCTCCGCCCAACT 146227

```

```

RESULT 8
AL353639/c DNA HTG 19-FEB-2001
LOCUS Homo sapiens chromosome 9 clone RP11-664D14, *** SEQUENCING IN
DEFINITION
ACCESSION
AL353639
VERSION AL353639.7 GI:13016506
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens

```

```

REFERENCE
AUTHORS Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE 1 (bases 1 to 196607)
JOURNAL Direct Submission
Submitted (18-FEB-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
Requests: clonerquests@sanger.ac.uk
On Feb 20, 2001 this sequence version replaced gi:12964414.

```

COMMENT ----- Genome Center

```

Center: Sanger Centre
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk
----- Project Information
Center project name: ba664D14
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Sequencing vector: plasmid; L08752; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Consensus quality: 195295 bases at least Q40
Consensus quality: 195779 bases at least Q20
Consensus quality: 196025 bases at least Q20
Insert size: 196207; sum-of-ctrls
Insert size: 186500; 3.2% error; agarose-fp
Quality coverage: 7.40x in Q20 bases; sum-of-ctrls Quality
coverage: 7.94x in Q20 bases; agarose-fp
-----

```

```

* NOTE: This is a 'working draft' sequence. It currently
* consists of 5 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

```

```

* 1 5394: contig of 5394 bp in length
* 5395 5494: gap of 100 bp
* 5495 21865: contig of 16371 bp in length
* 21866 21965: gap of 100 bp
* 21966 183547: contig of 161582 bp in length
* 183548 183647: gap of 100 bp
* 183648 194492: contig of 10845 bp in length
* 194493 194592: gap of 100 bp
* 194593 196607: contig of 2015 bp in length.

```

#### FEATURES

```

source
1..196607
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="9"
/clone="RP11-664D14"
/clone_11b="RPCI-11.3"

```











|            |   |  |             |
|------------|---|--|-------------|
| OY         | 4662  | ttctctcttgcatcttgtagaactaaaggaagaatcttgtagtgatccgctt             | 4661        |
| Db         | 279   | TTTTCTCTCTGCAATTTTGGAACCTTAAAGAAAGAAATATTTGTGTAAGGAATTCGCTT      | 220         |
| OY         | 4662  | ttactcttgaggaatctaaatgcttctgtagtcacatggccaagaattatcacgctaa       | 4721        |
| Db         | 219   | TTATTTTGTGGAAGATATCTAATGTTTGTGACTCACAATGGCCAAAGATTTATTACATGCTAAG | 160         |
| OY         | 4722  | ctggcttagtaaaagaagaataactctaaagctaaccaagaanaaibgccttcagtaagt     | 4781        |
| Db         | 159   | CTGGTTAGTATTAAGAAAGATATTTCTAAAGTTAACCAAGAAATGCTTCAGTAAAT         | 100         |
| OY         | 4782  | aggtatgaaaatgaaataataaaataagaagaanaatcccggggaggttcaaaaaaagc      | 4841        |
| Db         | 99  | AGGATGAAAAATGAAATATTAATAAGAAAGAAATTCGGGGAGTTTAAAAAAAATGC         | 40          |
| OY         | 4842  | ctcaattggcaatctaccctccctccccaaccaact                             | 4880        |
| Db         | 39  | CTCAATTTGGCAATCTACCTCTCTCCGCCACCCCAACT                           | 1           |
| RESULT     | 12  |  |             |
| LOCUS      | G27450/c  |  |             |
| DEFINITION | 339 bp DNA  | STS  | 28-JUN-1996 |
| ACCESSION  | human STS SHGC-32204, sequence tagged site.   |  |             |
| VERSION    | G27450  |  |             |
| KEYWORDS   | G27450.1 GI:1396169   |  |             |
| SOURCE     | STS; STS sequence; primer; sequence tagged site.  |  |             |
| ORGANISM   | human.  |  |             |
| REFERENCE  | Homo sapiens  |  |             |
| AUTHORS    | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;   |  |             |
| JOURNAL    | Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.   |  |             |
| COMMENT    | 1 (bases 1 to 339)<br>Myers,R.M.<br>Unpublished (1996)  |  |             |
|            | Contact: Richard M. Myers<br>Stanford Human Genome Center (SHGC)<br>Stanford University School of Medicine<br>Department of Genetics, M-344, Stanford, CA 94305, USA<br>Tel: 415/253687<br>Fax: 415/253689<br>Email: myers@shgc.stanford.edu  |  |             |
|            | Primer A: GGGGAGAGGAGTGAATGCG<br>Primer B: TCTAAGTTTAACCAAGAAATGCG<br>STS size: 125<br>PCR Profile:<br>Initial incubation: 94 degrees C for 90 seconds<br>Denaturation: 94 degrees C for 15 seconds<br>Annealing: 62 degrees C for 23 seconds<br>Polymerization: 72 degrees C for 30 seconds<br>PCR Cycles: 30<br>Thermal Cycler: Perkin Elmer 9600 |  |             |
|            | Protocol:<br>Template: 25 ng<br>Primer: each 1 uM<br>dNTPs: each 200 uM<br>Tag Polymerase: 0.05 units/uL<br>Total Vol: 10 uL  |  |             |
|            | Buffer:<br>MgCl2: 2.5 mM<br>KCl: 50 mM<br>Tris-HCl: 20 mM<br>pH: 8.3  |  |             |

Prepared with primer pairs provided by Sandoz, derived from Z39533  
-- Washington University/Merck EST sequence.  
Location/Qualifiers  
1..339

/organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /map="9"  
 11.135  
 primer\_bind 11.30  
 primer\_bind complement(111..135)  
 BASE COUNT 114 a 54 c 52 g 118 t 1 others  
 ORIGIN

Query Match 6.8%; Score 33.2; DB 54; Length 339;  
 Best Local Similarity 98.8%; Pred. No. 7.3e-71;  
 Matches 335; Conservative 0; Mismatches 4; Indels 0; Gaps 0

Oy 4542 taagttctttcttcacatagaatgcttctctactgacacacccaatcatcttctc 4601  
 Db 339 TAAgTTCTNTTTCATCTAGAAATGCTTCTTACTGACAACTTAACATCATTTT 280  
 Oy 4602 ttctctctgcatcttctggaacttaagaagaagaatatctgtcagtgcaatccgctt 4661  
 Db 279 TTTCTTCTTGGCACTTTTGTCACTTAAGGAAGGAATATTGTGATGGAATCTCGT 220  
 Oy 4662 ttattgtggaagatactcaatgtttctgagtcacatgagcagaatattacatgctaa 4721  
 Db 219 TTATTGTGGAGATATCTTAATTTTGTGATGTCACATGGCAAGAAATATTACATGCTAAG 160  
 Oy 4722 ctggttagtataagaagaataatctcaagctaacccaagaagaatgcttcagtaagt 4781  
 Db 159 CTGCTTACTATTAAGATAAGATTAATCTTAAGATTACCAAGAAATATGCTTCAGTAAAT 100  
 Oy 4782 aggaatgaagaatgaaatataataagaagaagaagaatctcgagggtttaaaaaaaatgc 4841  
 Db 99 AGGAAGAAAAAGAAAAATTAATAAATAAGAAAGAAATCTGGGGAGATTAAAAAAATGC 40  
 Oy 4842 ctcaattgcgcacatcactcctctcccccaccccaaat 4880  
 Db 39 CTCATTTGGCAATCTACTCTCTCTCCCAACCCCAACT 1

RESULT 13  
 AC013514 147739 bp DNA HTG 13-JUL-2000  
 LOCUS AC013514  
 DEFINITION Homo sapiens clone RP11-115J23, LOW-PASS SEQUENCE SAMPLING.  
 ACCESSION AC013514  
 VERSION AC013514.3 GI:9112399  
 KEYWORDS HTG; HTGS\_PHASEO.  
 SOURCE human.  
 ORGANISM Homo sapiens.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 REFERENCE 1 (bases 1 to 147739)  
 AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.  
 JOURNAL Homo sapiens, clone RP11-115J23  
 TITLE Unpublished  
 AUTHORS 2 (bases 1 to 147739)  
 Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,  
 Baldwin,J., Barna,N., Beckert,R., Boguslavsky,L., Boukhalter,B.,  
 Brown,A., Castle,A., Colangelo,M., Collins,S., Collymore,A.,  
 Cooke,P., DeRellano,K., Dewar,K., Domino,M., Donelan,L., Doyle,M.,  
 Ferreira,P., Fitzhugh,W., Forrest,C., Funke,R., Gage,D.,  
 Galagan,J., Gaidyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,  
 Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,  
 Lehoccky,J., Liew,C., Locke,K., MacDonald,P., Marquis,N.,  
 McMan,P., McGurt,A., McKernan,K., McLaughlin,J., Meldrum,J.,  
 Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,  
 Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,  
 Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,  
 Tesfaye,S., Tittrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,  
 Wyman,D., Ye,M.J., Zimmer,A. and Zody,M.  
 Direct Submission  
 Submitted (13-NOV-1999) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA  
 On Jul 13, 2000 this sequence version replaced gi:1653323.

All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center

Center: Whitehead Institute/MIT Center for Genome Research

Center code: MIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: [sequence\\_submissions@genome.wi.mit.edu](mailto:sequence_submissions@genome.wi.mit.edu)

----- Project Information

Center project name: L1888

Center clone name: 115\_J\_23

-----  
\* NOTE: This record contains 145 individual  
\* sequencing reads that have not been assembled into  
\* contigs. Runs of N are used to separate the reads  
\* and the order in which they appear is completely  
\* arbitrary. Low-pass sequence sampling is useful for  
\* identifying clones that may be gene-rich and allows  
\* overlap relationships among clones to be deduced.  
\* However, it should not be assumed that this clone  
\* will be sequenced to completion. In the event that  
\* the record is updated, the accession number will  
\* be preserved.

\* 1  
\* 901 1000: contig of 900 bp in length  
\* 1001 1936: contig of 936 bp in length  
\* 1937 2036: gap of 100 bp  
\* 2037 2954: contig of 918 bp in length  
\* 2955 3054: gap of 100 bp  
\* 3055 3990: contig of 936 bp in length  
\* 3991 4090: gap of 100 bp  
\* 4091 4990: contig of 900 bp in length  
\* 4991 5090: gap of 100 bp  
\* 5091 6035: contig of 945 bp in length  
\* 6036 6135: gap of 100 bp  
\* 6136 7056: contig of 921 bp in length  
\* 7057 7156: gap of 100 bp  
\* 7157 8072: contig of 916 bp in length  
\* 8073 8172: gap of 100 bp  
\* 8173 9111: contig of 939 bp in length  
\* 9112 9211: gap of 100 bp  
\* 9212 10132: contig of 921 bp in length  
\* 10133 10232: gap of 100 bp  
\* 10233 11162: contig of 930 bp in length  
\* 11163 12197: contig of 935 bp in length  
\* 12198 12297: gap of 100 bp  
\* 12298 13214: contig of 917 bp in length  
\* 13215 13314: gap of 100 bp  
\* 13315 14238: contig of 924 bp in length  
\* 14239 14338: gap of 100 bp  
\* 14339 15236: contig of 898 bp in length  
\* 15237 15336: gap of 100 bp  
\* 15337 16234: contig of 898 bp in length  
\* 16235 16334: gap of 100 bp  
\* 16335 17115: contig of 781 bp in length  
\* 17116 17215: gap of 100 bp  
\* 17216 18152: contig of 937 bp in length  
\* 18153 18252: gap of 100 bp  
\* 18253 19139: contig of 887 bp in length  
\* 19140 19239: gap of 100 bp  
\* 19240 20198: contig of 959 bp in length  
\* 20199 20298: gap of 100 bp  
\* 20299 21232: contig of 934 bp in length  
\* 21333 21332: gap of 100 bp  
\* 21333 22254: contig of 922 bp in length  
\* 22255 22354: gap of 100 bp  
\* 22355 23282: contig of 928 bp in length  
\* 23283 23382: gap of 100 bp  
\* 23383 24294: contig of 912 bp in length  
\* 24295 24394: gap of 100 bp  
\* 24395 25323: contig of 929 bp in length  
\* 25324 25423: gap of 100 bp

\* 25424 26378: contig of 955 bp in length  
\* 26379 26478: gap of 100 bp  
\* 26479 27442: contig of 964 bp in length  
\* 27443 27542: gap of 100 bp  
\* 27543 28472: contig of 930 bp in length  
\* 28473 28572: gap of 100 bp  
\* 28573 29512: contig of 940 bp in length  
\* 29513 29612: gap of 100 bp  
\* 29613 30519: contig of 907 bp in length  
\* 30520 30619: gap of 100 bp  
\* 30620 31561: contig of 942 bp in length  
\* 31562 31661: gap of 100 bp  
\* 31662 32590: contig of 929 bp in length  
\* 32591 32690: gap of 100 bp  
\* 32691 33629: contig of 939 bp in length  
\* 33630 33729: gap of 100 bp  
\* 33730 34659: contig of 930 bp in length  
\* 34660 34759: gap of 100 bp  
\* 34760 35664: contig of 905 bp in length  
\* 35665 35764: gap of 100 bp  
\* 35765 36711: contig of 947 bp in length  
\* 36712 36811: gap of 100 bp  
\* 36812 37755: contig of 944 bp in length  
\* 37756 37855: gap of 100 bp  
\* 37856 38799: contig of 944 bp in length  
\* 38800 38899: gap of 100 bp  
\* 38900 39813: contig of 914 bp in length  
\* 39814 39913: gap of 100 bp  
\* 39914 40822: contig of 909 bp in length  
\* 40823 40922: gap of 100 bp  
\* 40923 41787: contig of 865 bp in length  
\* 41788 41887: gap of 100 bp  
\* 41888 42801: contig of 914 bp in length  
\* 42802 42901: gap of 100 bp  
\* 42902 43816: contig of 915 bp in length  
\* 43817 43916: gap of 100 bp  
\* 43917 44843: contig of 927 bp in length  
\* 44844 44943: gap of 100 bp  
\* 44944 45868: contig of 925 bp in length  
\* 45869 45968: gap of 100 bp  
\* 45969 46900: contig of 932 bp in length  
\* 46901 47000: gap of 100 bp  
\* 47001 47910: contig of 910 bp in length  
\* 47911 48010: gap of 100 bp  
\* 48011 48951: contig of 941 bp in length  
\* 48952 49051: gap of 100 bp  
\* 49052 50001: contig of 950 bp in length  
\* 50002 50101: gap of 100 bp  
\* 50102 51030: contig of 929 bp in length  
\* 51031 51130: gap of 100 bp  
\* 51131 52073: contig of 943 bp in length  
\* 52074 52173: gap of 100 bp  
\* 52174 53096: contig of 923 bp in length  
\* 53097 53196: gap of 100 bp  
\* 53197 54128: contig of 932 bp in length  
\* 54129 54228: gap of 100 bp  
\* 54229 55147: contig of 919 bp in length  
\* 55148 55247: gap of 100 bp  
\* 55248 56175: contig of 928 bp in length  
\* 56176 56275: gap of 100 bp  
\* 56276 57210: contig of 935 bp in length  
\* 57211 57310: gap of 100 bp  
\* 57311 58235: contig of 925 bp in length  
\* 58236 58335: gap of 100 bp  
\* 58336 59284: contig of 949 bp in length  
\* 59285 59384: gap of 100 bp  
\* 59385 60309: contig of 925 bp in length  
\* 60310 60409: gap of 100 bp  
\* 60410 61340: contig of 931 bp in length  
\* 61341 61440: gap of 100 bp  
\* 61441 62399: contig of 959 bp in length  
\* 62400 62499: gap of 100 bp  
\* 62500 63435: contig of 936 bp in length

NOTES: BATTEN, B.; ELLISON, D.; NUSBAUM, C. and LAUNDER, E.

|   |       |        |           |        |           |
|---|-------|--------|-----------|--------|-----------|
| * | 1     | 88:    | contlg of | 898 bp | in length |
| * | 899   | 998:   | gap of    | 100 bp |           |
| * | 999   | 1871:  | contlg of | 873 bp | in length |
| * | 1872  | 1971:  | gap of    | 100 bp |           |
| * | 1972  | 2855:  | contlg of | 874 bp | in length |
| * | 2846  | 2945:  | gap of    | 100 bp |           |
| * | 2946  | 3797:  | contlg of | 852 bp | in length |
| * | 3798  | 3847:  | gap of    | 100 bp |           |
| * | 3898  | 4782:  | contlg of | 885 bp | in length |
| * | 4783  | 4882:  | gap of    | 100 bp |           |
| * | 4883  | 5767:  | contlg of | 885 bp | in length |
| * | 5768  | 5867:  | gap of    | 100 bp |           |
| * | 5868  | 6777:  | contlg of | 910 bp | in length |
| * | 6778  | 6877:  | gap of    | 100 bp |           |
| * | 6878  | 7751:  | contlg of | 874 bp | in length |
| * | 7752  | 7851:  | gap of    | 100 bp |           |
| * | 7852  | 8760:  | contlg of | 909 bp | in length |
| * | 8761  | 8860:  | gap of    | 100 bp |           |
| * | 8861  | 9757:  | contlg of | 877 bp | in length |
| * | 9758  | 9837:  | gap of    | 100 bp |           |
| * | 9838  | 10713: | contlg of | 876 bp | in length |
| * | 10714 | 10813: | gap of    | 100 bp |           |
| * | 10814 | 11772: | contlg of | 909 bp | in length |
| * | 11723 | 11822: | gap of    | 100 bp |           |
| * | 11823 | 12706: | contlg of | 884 bp | in length |
| * | 12707 | 12806: | gap of    | 100 bp |           |
| * | 12807 | 13680: | contlg of | 874 bp | in length |
| * | 13681 | 13780: | gap of    | 100 bp |           |

13781 14684: contig of 904 bp in length  
14685 14784: gap of 100 bp  
14785 15678: contig of 894 bp in length  
15679 15778: gap of 100 bp  
15779 16663: contig of 885 bp in length  
16664 16763: gap of 100 bp  
16764 17652: contig of 889 bp in length  
17653 17752: gap of 100 bp  
17753 18642: contig of 890 bp in length  
18643 18742: gap of 100 bp  
18743 19647: contig of 905 bp in length  
19648 19747: gap of 100 bp  
19748 20663: contig of 916 bp in length  
20664 20763: gap of 100 bp  
20764 21656: contig of 893 bp in length  
21657 21756: gap of 100 bp  
21757 22670: contig of 914 bp in length  
22671 22770: gap of 100 bp  
22771 23662: contig of 892 bp in length  
23663 23762: gap of 100 bp  
23763 24661: contig of 899 bp in length  
24662 24761: gap of 100 bp  
24762 25666: contig of 905 bp in length  
25667 25766: gap of 100 bp  
25767 26649: contig of 883 bp in length  
26650 26749: gap of 100 bp  
26750 27632: contig of 883 bp in length  
27633 27732: gap of 100 bp  
27733 28592: contig of 860 bp in length  
28593 28692: gap of 100 bp  
28693 29394: contig of 902 bp in length  
29395 29694: gap of 100 bp  
29695 30572: contig of 878 bp in length  
30573 30672: gap of 100 bp  
30673 31558: contig of 887 bp in length  
31559 31659: gap of 100 bp  
31660 32553: contig of 894 bp in length  
32554 32653: gap of 100 bp  
32654 33547: contig of 894 bp in length  
33548 33647: gap of 100 bp  
33648 34546: contig of 899 bp in length  
34547 34646: gap of 100 bp  
34647 35531: contig of 885 bp in length  
35532 35631: gap of 100 bp  
35632 36515: contig of 884 bp in length  
36516 36615: gap of 100 bp  
36616 37509: contig of 894 bp in length  
37510 37609: gap of 100 bp  
37609 38508: contig of 899 bp in length  
38509 38608: gap of 100 bp  
38609 39489: contig of 881 bp in length  
39490 39589: gap of 100 bp  
39590 40474: contig of 885 bp in length  
40475 40574: gap of 100 bp  
40575 41466: contig of 892 bp in length  
41467 41566: gap of 100 bp  
41567 42495: contig of 929 bp in length  
42496 42595: gap of 100 bp  
42596 43486: contig of 891 bp in length  
43487 43586: gap of 100 bp  
43587 44470: contig of 884 bp in length  
44471 44570: gap of 100 bp  
44571 45461: contig of 891 bp in length  
45462 45561: gap of 100 bp  
45562 46458: contig of 897 bp in length  
46459 46558: gap of 100 bp  
46559 47444: contig of 886 bp in length  
47445 47544: gap of 100 bp  
47545 48425: contig of 881 bp in length  
48426 48525: gap of 100 bp  
48526 49407: contig of 882 bp in length  
49408 49507: gap of 100 bp  
49508 50412: contig of 905 bp in length

50413 50512: gap of 100 bp  
50513 51417: contig of 905 bp in length  
51418 51517: gap of 100 bp  
51518 52413: contig of 896 bp in length  
52414 52513: gap of 100 bp  
52514 53414: contig of 901 bp in length  
53415 53514: gap of 100 bp  
53515 54425: contig of 911 bp in length  
54426 54525: gap of 100 bp  
54526 55423: contig of 898 bp in length  
55424 55523: gap of 100 bp  
55524 56431: contig of 908 bp in length  
56432 56531: gap of 100 bp  
56532 57422: contig of 891 bp in length  
57423 57522: gap of 100 bp  
57523 58426: contig of 904 bp in length  
58427 58526: gap of 100 bp  
58527 59435: contig of 909 bp in length  
59436 59535: gap of 100 bp  
59536 60438: contig of 903 bp in length  
60439 60538: gap of 100 bp  
60539 61434: contig of 896 bp in length  
61435 61534: gap of 100 bp  
61535 62423: contig of 889 bp in length  
62424 62523: gap of 100 bp  
62524 63400: contig of 877 bp in length  
63401 63500: gap of 100 bp  
63501 64400: contig of 900 bp in length  
64401 64500: gap of 100 bp  
64501 65399: contig of 899 bp in length  
65400 65499: gap of 100 bp  
65500 66385: contig of 886 bp in length  
66386 66485: gap of 100 bp  
66486 67358: contig of 873 bp in length  
67359 67458: gap of 100 bp  
67459 68359: contig of 901 bp in length  
68360 68459: gap of 100 bp  
68460 69340: contig of 881 bp in length  
69341 69440: gap of 100 bp  
69441 70326: contig of 886 bp in length  
70327 70426: gap of 100 bp  
70427 71332: contig of 906 bp in length  
71333 71432: gap of 100 bp

Query Match  
Best Local Similarity 99.08; Score 285.2; DB 61; Length 149824;  
Matches 287; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1058 agagtgatgtaaaatctctgtgataccagtaactgatactgacatgctcctggaggaac 1117  
|||||  
Db 60865 AGGATGTTATGAAATTCCTGTCATCCAGTACTGATCGCATATGCTCTCGAGGAC 60806  
QY 1118 tataaccagaatctctcctgaaagacagatggaataacacctggtagacataa 1177  
|||||  
Db 60805 TATATACCAAGATCTCTCGCAAGACAGATGAGATATACCTTCGCTGACATAAGTA 60746  
QY 1178 tggggcctgctctcctgcttactaataatgactaacacacctgcaaatgctatgaaac 1237  
|||||  
Db 60745 TGGGGCCTGCTTGGCTTACTATAATGACTACACACCTGCAAAATGATTGAACAC 60686  
QY 1238 aatatgaatgtaaaacacataatgtagtgaactgaatcattcaatcagaagtatat 1297  
|||||  
Db 60685 AATATGAATGTGAAACACATAGTGGACTGATTCATTACCAAGTAAGTATATAT 60626  
QY 1298 caatgagagactcctcctgcttaccagatcagctggaaggagctc 1347  
|||||  
Db 60625 CAAGTCAGAACTTCTTCTGTGCTACCCGATTCAGCTGGAAGGAGATCT 60576

RESULT 15  
AC013514/c  
LOCUS  
DEFINITION Homo sapiens clone RP11-115J23, LOW-PASS SEQUENCE SAMPLING.  
AC013514 147739 bp DNA HTG 13-JUL-2000

ACCESSION AC013514  
VERSION AC013514.3 GI:9112399  
KEYWORDS HTG, HTGS\_PHASE0.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 147739)  
AUTHORS Birren, B., Linton, L., Nusbaum, C. and Lander, E.  
TITLE Homo sapiens, clone RP11-115723  
JOURNAL Unpublished  
AUTHORS 2 (bases 1 to 147739)  
Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, M.,  
Baldwin, J., Barna, N., Beckert, R., Boguslavskiy, L., Bouckgalter, B.,  
Brown, A., Castle, A., Colangelo, M., Collins, S., Collinmore, A.,  
Cooke, P., Dearlano, K., Dewar, K., Domino, M., Donelan, L., Doyle, M.,  
Ferreira, P., FitzHugh, W., Forrest, C., Funke, R., Gage, D.,  
Galagan, J., Gardyna, S., Grant, G., Hagos, B., Heaford, A., Horton, L.,  
Howland, J., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J.,  
Lehocky, J., Lieu, C., Locke, K., Macdonald, P., Marquis, N.,  
McKen, P., McGurk, A., McKernan, K., McLaughlin, J., Meldrum, J.,  
Morrow, J., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P.,  
Peterson, K., Pollara, V., Riley, R., Roy, A., Santos, R., Severy, P.,  
Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,  
Tesfaye, S., Tirrell, A., Vassiliev, H., Vo, A., Wheeler, J., Wu, X.,  
Wyman, D., Ye, W. J., Zimmer, A. and Zody, M.  
Direct Submission  
Submitted (13-NOV-1999) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA.  
On Jul 13, 2000 this sequence version replaced gi.16533323.  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>  
----- Genome Center  
Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIBR  
Web site: <http://www-seq.wi.mit.edu>  
Contact: sequence.submissions@genome.wi.mit.edu  
----- Project Information  
Center project name: L1888  
Center clone name: 115\_J\_23  
-----  
\* NOTE: This record contains 145 individual  
\* sequencing reads that have not been assembled into  
\* contigs. Runs of N are used to separate the reads  
\* and the order in which they appear is completely  
\* arbitrary. Low-pass sequence sampling is useful for  
\* identifying clones that may be gene-rich and allows  
\* overlap relationships among clones to be deduced.  
\* However, it should not be used to deduce that this clone  
\* will be sequenced to completion. In the event that  
\* the record is updated, the accession number will  
\* be preserved.  
\* 1  
\* 900: contig of 900 bp in length  
\* 901 1000: gap of 100 bp  
\* 1001 1936: contig of 936 bp in length  
\* 1937 2036: gap of 100 bp  
\* 2037 2954: contig of 918 bp in length  
\* 2955 3054: gap of 100 bp  
\* 3055 3990: contig of 936 bp in length  
\* 3991 4090: gap of 100 bp  
\* 4091 4990: contig of 900 bp in length  
\* 4991 5090: gap of 100 bp  
\* 5091 6035: contig of 945 bp in length  
\* 6036 6135: gap of 100 bp  
\* 6136 7056: contig of 921 bp in length  
\* 7057 7156: gap of 100 bp  
\* 7157 8072: contig of 916 bp in length  
\* 8073 8172: gap of 100 bp  
\* 8173 9111: contig of 938 bp in length  
\* 9112 9211: gap of 100 bp  
\* 9212 10132: contig of 921 bp in length  
\* 10133 10232: gap of 100 bp

10233 11162: contig of 930 bp in length  
11163 11262: gap of 100 bp  
11263 12197: contig of 935 bp in length  
12198 12297: gap of 100 bp  
12298 13214: contig of 917 bp in length  
13215 13314: gap of 100 bp  
13315 14238: contig of 924 bp in length  
14239 14338: gap of 100 bp  
14339 15236: contig of 898 bp in length  
15237 15336: gap of 100 bp  
15337 16234: contig of 898 bp in length  
16235 16334: gap of 100 bp  
16335 17115: contig of 781 bp in length  
17116 17215: gap of 100 bp  
17216 18152: contig of 937 bp in length  
18153 18252: gap of 100 bp  
18253 19139: contig of 887 bp in length  
19140 19239: gap of 100 bp  
19240 20198: contig of 959 bp in length  
20199 20298: gap of 100 bp  
20299 21232: contig of 934 bp in length  
21233 21332: gap of 100 bp  
21333 22254: contig of 922 bp in length  
22255 22354: gap of 100 bp  
22355 23282: contig of 928 bp in length  
23283 23382: gap of 100 bp  
23383 24294: contig of 912 bp in length  
24295 24394: gap of 100 bp  
24395 25323: contig of 929 bp in length  
25324 25423: gap of 100 bp  
25424 26378: contig of 955 bp in length  
26379 26478: gap of 100 bp  
26479 27442: contig of 964 bp in length  
27443 27542: gap of 100 bp  
27543 28472: contig of 930 bp in length  
28473 28572: gap of 100 bp  
28573 29512: contig of 940 bp in length  
29513 29612: gap of 100 bp  
29613 30519: contig of 907 bp in length  
30520 30619: gap of 100 bp  
30620 31561: contig of 942 bp in length  
31562 31661: gap of 100 bp  
31662 32590: contig of 929 bp in length  
32591 32690: gap of 100 bp  
32691 33623: contig of 939 bp in length  
33630 33729: gap of 100 bp  
33730 34659: contig of 930 bp in length  
34660 34759: gap of 100 bp  
34760 35664: contig of 905 bp in length  
35665 35764: gap of 100 bp  
35765 36711: contig of 947 bp in length  
36712 36811: gap of 100 bp  
36812 37755: contig of 944 bp in length  
37756 37855: gap of 100 bp  
37856 38799: contig of 944 bp in length  
38800 38899: gap of 100 bp  
38900 39813: contig of 914 bp in length  
39814 39913: gap of 100 bp  
39914 40822: contig of 909 bp in length  
40823 40922: gap of 100 bp  
40923 41787: contig of 865 bp in length  
41788 41887: gap of 100 bp  
41888 42801: contig of 914 bp in length  
42802 42901: gap of 100 bp  
42902 43816: contig of 915 bp in length  
43817 43916: gap of 100 bp  
43917 44843: contig of 927 bp in length  
44844 44943: gap of 100 bp  
44944 45868: contig of 925 bp in length  
45869 45968: gap of 100 bp  
45969 46900: contig of 932 bp in length  
46901 47000: gap of 100 bp  
47001 47910: contig of 910 bp in length









GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 12, 2001, 18:58:57 ; Search time 284.56 Seconds

(without alignments)  
10768.063 Million cell updates/sec

Title: US-09-502-698-3

Perfect score: 4880  
Sequence: 1 cccggccggcgacagtg.....tctctcccaacccaact 4880

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 730101 seqs, 313950809 residues

Total number of hits satisfying chosen parameters: 1460202

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq.0601.\*

|     |  |
|-----|--|
| 1:  | /SIDSI/gcgdata/geneseq/geneseqn/NA1980.DAT.* |
| 2:  | /SIDSI/gcgdata/geneseq/geneseqn/NA1981.DAT.* |
| 3:  | /SIDSI/gcgdata/geneseq/geneseqn/NA1982.DAT.* |
| 4:  | /SIDSI/gcgdata/geneseq/geneseqn/NA1983.DAT.* |
| 5:  | /SIDSI/gcgdata/geneseq/geneseqn/NA1984.DAT.* |
| 6:  | /SIDSI/gcgdata/geneseq/geneseqn/NA1985.DAT.* |
| 7:  | /SIDSI/gcgdata/geneseq/geneseqn/NA1986.DAT.* |
| 8:  | /SIDSI/gcgdata/geneseq/geneseqn/NA1987.DAT.* |
| 9:  | /SIDSI/gcgdata/geneseq/geneseqn/NA1988.DAT.* |
| 10: | /SIDSI/gcgdata/geneseq/geneseqn/NA1989.DAT.* |
| 11: | /SIDSI/gcgdata/geneseq/geneseqn/NA1990.DAT.* |
| 12: | /SIDSI/gcgdata/geneseq/geneseqn/NA1991.DAT.* |
| 13: | /SIDSI/gcgdata/geneseq/geneseqn/NA1992.DAT.* |
| 14: | /SIDSI/gcgdata/geneseq/geneseqn/NA1993.DAT.* |
| 15: | /SIDSI/gcgdata/geneseq/geneseqn/NA1994.DAT.* |
| 16: | /SIDSI/gcgdata/geneseq/geneseqn/NA1995.DAT.* |
| 17: | /SIDSI/gcgdata/geneseq/geneseqn/NA1996.DAT.* |
| 18: | /SIDSI/gcgdata/geneseq/geneseqn/NA1997.DAT.* |
| 19: | /SIDSI/gcgdata/geneseq/geneseqn/NA1998.DAT.* |
| 20: | /SIDSI/gcgdata/geneseq/geneseqn/NA1999.DAT.* |
| 21: | /SIDSI/gcgdata/geneseq/geneseqn/NA2000.DAT.* |
| 22: | /SIDSI/gcgdata/geneseq/geneseqn/NA2001.DAT.* |

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score  | Query Match | Length | DB ID | Description |
|------------|--------|-------------|--------|-------|-------------|
| 1          | 4880   | 100.0       | 4880   | 20    | AAx29908    |
| 2          | 4662.6 | 95.5        | 6540   | 20    | AAx29909    |
| 3          | 4525.6 | 92.7        | 7431   | 21    | AAx236453   |
| 4          | 4436.2 | 90.9        | 6750   | 20    | AAx29910    |
| 5          | 2825   | 57.9        | 3188   | 21    | AAx47424    |
| 6          | 2781.8 | 57.0        | 2819   | 20    | AAx29960    |
| 7          | 2722.2 | 55.8        | 6666   | 20    | AAx29911    |
| 8          | 1433.4 | 29.4        | 1681   | 20    | AAx52869    |
| 9          | 854.4  | 17.5        | 2067   | 19    | AAV69613    |
| 10         | 753.4  | 15.4        | 755    | 20    | AAx29969    |
| 11         | 730.4  | 15.0        | 865    | 20    | AAx29970    |

|    |       |      |      |    |          |  |
|----|-------|------|------|----|----------|--|
| 12 | 730.4 | 15.0 | 959  | 20 | AAx29971 | PDZ domain-contains PDZ domain-contains DNA encoding the h |
| 13 | 563.6 | 11.5 | 1776 | 20 | AAx29967 | DNA encoding the h   |
| 14 | 540.2 | 11.1 | 5836 | 20 | AAx86366 | Human secreted pro   |
| 15 | 124   | 2.5  | 229  | 21 | AAx27886 | Human biallelic po   |
| 16 | 123   | 2.5  | 125  | 19 | AAx1808  | Human biallelic po   |
| 17 | 123   | 2.5  | 125  | 19 | AAx12075 | Human biallelic po   |
| 18 | 122   | 2.5  | 361  | 20 | AAV88388 | EST clone GP107.   |
| 19 | 121.4 | 2.5  | 125  | 19 | AAx1809  | Human biallelic po   |
| 20 | 108.6 | 2.2  | 936  | 22 | AAx58252 | Oligonucleotide D1   |
| 21 | 108.6 | 2.2  | 936  | 22 | AAx58254 | Oligonucleotide D1   |
| 22 | 108.6 | 2.2  | 936  | 22 | AAx58257 | Oligonucleotide D1   |
| 23 | 108.6 | 2.2  | 936  | 22 | AAx58259 | Oligonucleotide D2   |
| 24 | 108.6 | 2.2  | 936  | 22 | AAx58262 | Oligonucleotide D2   |
| 25 | 108.6 | 2.2  | 938  | 22 | AAx58255 | Oligonucleotide D1   |
| 26 | 103.4 | 2.1  | 936  | 22 | AAx58252 | Oligonucleotide D1   |
| 27 | 103.4 | 2.1  | 936  | 22 | AAx58254 | Oligonucleotide D1   |
| 28 | 103.4 | 2.1  | 936  | 22 | AAx58257 | Oligonucleotide D1   |
| 29 | 103.4 | 2.1  | 936  | 22 | AAx58259 | Oligonucleotide D1   |
| 30 | 103.4 | 2.1  | 936  | 22 | AAx58262 | Oligonucleotide D2   |
| 31 | 103.4 | 2.1  | 938  | 22 | AAx58255 | Oligonucleotide D1   |
| 32 | 79.2  | 1.6  | 214  | 21 | AAx03452 | Human secreted pro   |
| 33 | 66    | 1.4  | 8043 | 16 | AAx05924 | Human protein tyro   |
| 34 | 66    | 1.4  | 8119 | 21 | AAx61840 | DNA encoding a hum   |
| 35 | 65.4  | 1.3  | 2963 | 18 | AAx93775 | PSD-93 coding sequ   |
| 36 | 63.6  | 1.3  | 1000 | 17 | AAx18384 | Human Fas-associat   |
| 37 | 63.6  | 1.3  | 1830 | 17 | AAx18383 | Human Fas-associat   |
| 38 | 63    | 1.3  | 2534 | 21 | AAx01182 | PDZ domain-contains PDZ domain-contains                    |
| 39 | 62    | 1.3  | 1060 | 21 | AAx98033 | Human colon cancer   |
| 40 | 62    | 1.3  | 2249 | 22 | AAx24178 | Human secreted pro   |
| 41 | 62    | 1.3  | 2308 | 21 | AAx47426 | Sequence encoding  |
| 42 | 62    | 1.3  | 2571 | 22 | AAx5028  | Membrane-bound pro   |
| 43 | 62    | 1.3  | 2571 | 22 | AAx92077 | Human PRO1136 CDNA   |
| 44 | 62    | 1.3  | 2571 | 22 | AAx44174 | Human PRO1136 (UNQ   |
| 45 | 62    | 1.3  | 5367 | 22 | AAx24155 | Human secreted pro   |

## ALIGNMENTS

|          |   |                         |  |  |  |
|----------|---|-------------------------|--|--|--|
| RESULT 1 |   |                         |  |  |  |
| ID       | AAx29908  | standard; DNA; 4880 BP. |  |  |  |
| AC       | AAx29908;   |                         |  |  |  |
| XX       |   |                         |  |  |  |
| DT       | 06-JUL-1999   | (first entry)           |  |  |  |
| XX       |   |                         |  |  |  |
| DE       | Clone 38-2-1 encoding protein containing PDZ domain.  |                         |  |  |  |
| XX       |   |                         |  |  |  |
| KW       | PDZ domain; gene expression; human umbilical vascular endothelial cell;   |                         |  |  |  |
| KW       | HUVCC; stimulation; tumour necrosis factor; TNF; protein binding;   |                         |  |  |  |
| KW       | cell; proliferation disorder; cancer; ss.   |                         |  |  |  |
| XX       |   |                         |  |  |  |
| OS       | Homo sapiens.   |                         |  |  |  |
| XX       |   |                         |  |  |  |
| PN       | W09907846-A1.   |                         |  |  |  |
| XX       |   |                         |  |  |  |
| PD       | 18-FEB-1999.  |                         |  |  |  |
| XX       |   |                         |  |  |  |
| PF       | 12-AUG-1998;  | 98WO-JP03603.           |  |  |  |
| XX       |   |                         |  |  |  |
| PR       | 19-JUN-1998;  | 98JP-0189944.           |  |  |  |
| XX       |   |                         |  |  |  |
| PR       | 12-AUG-1997;  | 97JP-0230356.           |  |  |  |
| XX       |   |                         |  |  |  |
| PA       | (CHUG-) CHUGAI RES INST MOLECULAR MEDICINE INC.   |                         |  |  |  |
| XX       |   |                         |  |  |  |
| PI       | Funahashi S, Miyata S;  |                         |  |  |  |
| XX       |   |                         |  |  |  |
| DR       | WPI: 1999-167423/14.  |                         |  |  |  |
| XX       |   |                         |  |  |  |
| DR       | P-PSDB; AA104730, AA104731.   |                         |  |  |  |
| XX       |   |                         |  |  |  |
| PT       | Protein containing PDZ domain, whose expression is enhanced by TNF stimulation - plays an important role in protein/protein |                         |  |  |  |



|    |      |   |      |
|----|------|---|------|
| QY | 3001 | atccgaataacagcagatcacaacacagcaattttgtctctgatactctgacaactcgc           | 3060 |
| Db | 3001 | atccgaataataagcagatcatcatcaacaacagcaattttgtctctgatactctgacaactcgc     | 3060 |
| QY | 3061 | cccatcttcctctgctctgcgaaacaaacatcgcgattttccaaggcgcgaacgcgctggc         | 3120 |
| Db | 3061 | cccatcttcctctgctctgcgaaacaaacatcgcgattttccaaggcgcgaaacgcgctggc        | 3120 |
| QY | 3121 | ctgagcactcgtctgggggttcagacacgcctcgccttgagccttatatccagaagtttat         | 3180 |
| Db | 3121 | ctgagcactcgtctgggggttcagacacgcctcgccttgagccttatatccagaagtttat         | 3180 |
| QY | 3181 | gaagaagagcagcagatctataaagaatggaagactctggcctggaatcagatcttagagctg       | 3240 |
| Db | 3181 | gaagaagagcagcagatctataaagaatggaagactctggcctggaatcagatcttagagctg       | 3240 |
| QY | 3241 | aatggaattgacttgaggaagcgcacacatgattgaagcaataatgtctctgagacagcg          | 3300 |
| Db | 3241 | aatggaattgacttgaggaagcgcacacatgattgaagcaataatgtctctgagacagcg          | 3300 |
| QY | 3301 | ccacagagagctgcgcctgcgaacactctacaagatctgaagccccaataaagagaggaagctg      | 3360 |
| Db | 3301 | ccacagagagctgcgcctgcgaacactctacaagatctgaagccccaataaagagaggaagctg      | 3360 |
| QY | 3361 | ctgacacccctcaactatctgaactgcgaagaagccgcggaaaaagccttagatatt             | 3420 |
| Db | 3361 | ctgacacccctcaactatctgaactgcgaagaagccgcggaaaaagccttagatatt             | 3420 |
| QY | 3421 | gttggtaaaagaaacgatactagagatatttgctcagacattgtctaaaggcggaattgcga        | 3480 |
| Db | 3421 | gttggtaaaagaaacgatactagagatatttgctcagacattgtctaaaggcggaattgcga        | 3480 |
| QY | 3481 | gatacccgatggaagaacgatactcgaaggagacagatacttatctgtgtaagcgggaagacgctt    | 3540 |
| Db | 3481 | gatacccgatggaagaacgatactcgaaggagacagatacttatctgtgtaagcgggaagacgctt    | 3540 |
| QY | 3541 | cgtaatgcctctccaaagaagcggcttgccgcctctgtctaagtgcttccttagcacagtaacc      | 3600 |
| Db | 3541 | cgtaatgcctctccaaagaagcggcttgccgcctctgtctaagtgcttccttagcacagtaacc      | 3600 |
| QY | 3601 | ctggaagtctggaagaatcacaagcttgcttcacatctcaagagagagagcattcccaacc         | 3660 |
| Db | 3601 | ctggaagtctggaagaatcacaagcttgcttcacatctcaagagagagagcattcccaacc         | 3660 |
| QY | 3661 | agccaagctgagctggaagcagacgcctctctcttcacatttcacactctcgcgataccaagata     | 3720 |
| Db | 3661 | agccaagctgagctggaagcagacgcctctctctcttcacatttcacactctcgcgataccaagata   | 3720 |
| QY | 3721 | ctctgatactactggaagaatgactctcaaaagaagaatctgcgaatctctgaataacaggaatta    | 3780 |
| Db | 3721 | ctctgatactactggaagaatgactctcaaaagaagaatctgcgaatctctgaataacaggaatta    | 3780 |
| QY | 3781 | agaaacagctcgaaatcgaaaagggccctactgactcaactctgggaatacagatactgcctggagga  | 3840 |
| Db | 3781 | agaaacagctcgaaatcgaaaagggccctactgactcaactctgggaatacagatactgcctggagga  | 3840 |
| QY | 3841 | gtaagcaagcccaactctgtgtgtgtccctataattatctgaaatctgacccaacgcggagctt      | 3900 |
| Db | 3841 | gtaagcaagcccaactctgtgtgtgtccctataattatctgaaatctgacccaacgcggagctt      | 3900 |
| QY | 3901 | gcagcacaagaccccaaaaacccaagattctggggaatagatctgtcacacatctgtgtgcacatcc   | 3960 |
| Db | 3901 | gcagcacaagaccccaaaaacccaagattctggggaatagatctgtcacacatctgtgtgcacatcc   | 3960 |
| QY | 3961 | actgagggcactgactcacaacccaagcagcttaacactctctgaaaaatgtcatctcgctccatct   | 4020 |
| Db | 3961 | actgagggcactgactcacaacccaagcagcttaacactctctgaaaaatgtcatctcgctccatct   | 4020 |
| QY | 4021 | gaatctgagcagctggtctgtctggaaggaagcgtgagctgtgtctacaagctgcacatccagagcct  | 4080 |
| Db | 4021 | gaatctgagcagctggtctgtctggaaggaagcgtgagctgtgtctacaagctgcacatccagagcct  | 4080 |
| QY | 4081 | gcgaagcttcacgctcttcttcacatctggcctgcagcctcaaccacagatattctcagagatgattta | 4140 |

|  |      |   |      |
|--|------|---|------|
| Db   | 4081 | ggaagctccagctcttcttccactcgtggcgcgaagctcaacagatatttcaagtgattta       | 4140 |
| Qy   | 4141 | ggaactctccaatgttaagctctattacacacacagagcgagagccagatggtcttagctcagt    | 4200 |
| Db   | 4141 | ggaccctcccaatgttaagctctattacacacagagcgagagccagatggtcttagctcagt      | 4200 |
| Qy   | 4201 | atgagttggaggatataagcagccctcaatggaaagcttcaacattatgtttaaacaagtgctt    | 4260 |
| Db   | 4201 | atgagttggaggatataagcagccctcaatggaaagcttcaacattatgtttaaacaagtgctt    | 4260 |
| Qy   | 4261 | gcaaaaggagcagcgccctctgaagacggaagctctgtaaaaaggcgatcagatctgtctgc      | 4320 |
| Db   | 4261 | gcaaaaggagcagcgccctctgaagacggaagctctgtaaaaaggcgatcagatctgtctgc      | 4320 |
| Qy   | 4321 | aatggcgagagctctgaagaagtgtaecccattgaagaagcggttgccattccttaacggaca     | 4380 |
| Db   | 4321 | aatggcgagagctctgaagaagtgtaecccattgaagaagcggttgccattccttaacggaca     | 4380 |
| Qy   | 4381 | aaaggaactgtaactgttgattgtgtctctctcttaattatggtgcgcagaaattgaaacaccca   | 4440 |
| Db   | 4381 | aaaggaactgtaactgttgattgtgtctctctctcttaattatggtgcgcagaaattgaaacaccca | 4440 |
| Qy   | 4441 | accctatgctcaactcctctactctgtlaaagagaatgcaactggtcctgcacaattttatgctgt  | 4500 |
| Db   | 4441 | accctatgctcaactcctctactctgtlaaagagaatgcaactggtcctgcacaattttatgctgt  | 4500 |
| Qy   | 4501 | gtctcagccgggtctctcaaaactgttaggggggaaataacattaaattctctttctcacc       | 4560 |
| Db   | 4501 | gtctcagccgggtctctcaaaactgttaggggggaaataacattaaattctctttctcacc       | 4560 |
| Qy   | 4561 | tggaaatgctcttccttaactgtgacaacccaatcaatctttctcttctccttgcaatttggtg    | 4620 |
| Db   | 4561 | tggaaatgctcttccttaactgtgacaacccaatcaatctttctcttctccttgcaatttggtg    | 4620 |
| Qy   | 4621 | aacttaaaagaagaagatattgtgtgagtgaaatcgtttttatttttgaggagatatcta        | 4680 |
| Db   | 4621 | aacttaaaagaagaagatattgtgtgagtgaaatcgtttttatttttgaggagatatcta        | 4680 |
| Qy   | 4681 | atgcttttgagtcacacatgvgcagaagaattctcaatgcttaagctggtgtatataagaagaag   | 4740 |
| Db   | 4681 | atgcttttgagtcacacatgvgcagaagaattctcaatgcttaagctggtgtatataagaagaag   | 4740 |
| Qy   | 4741 | ataattcttaagcttaaccaaagaagaattgcttcagtaagtttggatgaaaaatgaaataa      | 4800 |
| Db   | 4741 | ataattcttaagcttaaccaaagaagaattgcttcagtaagtttggatgaaaaatgaaataa      | 4800 |
| Qy   | 4801 | taaaaataaagaagaagaatctcgvggagttctaaaaaaaatgacctcaattgvgcactacc      | 4860 |
| Db   | 4801 | taaaaataaagaagaagaatctcgvggagttctaaaaaaaatgacctcaattgvgcactacc      | 4860 |
| Qy   | 4861 | tccctctccaccaccaact 4880  |      |
| Db   | 4861 | tccctctccaccaccaact 4880  |      |
| resulr 2   |      |   |      |
| AAK29909   |      |   |      |
| ID AAK29909 standard; DNA; 6540 BP.  |      |   |      |
| XX AAK29909;   |      |   |      |
| XX DT 06-JUL-1999 (first entry)  |      |   |      |
| XX Clone 38-2-1a encoding protein containing PDZ domain.                   |      |   |      |
| XX PDZ domain; gene expression; human umbilical vascular endothelial cell; |      |   |      |
| XX HUVEC; stimulation; tumour necrosis factor; TNF; protein binding;       |      |   |      |
| XX cell; proliferation disorder; cancer; ss.                               |      |   |      |
| OS Homo sapiens.   |      |   |      |
| XX   |      |   |      |
| PN W09907846-A1.   |      |   |      |

[illegible]

[illegible]

|    |      |  |      |
|----|------|--|------|
| Dh | 3492 | cagacataagacaattggaaatcaagcccacagggggttggaactctcgagagaaacaaagacaat     | 3551 |
| Qy | 1892 | ccttaagacatcaagatctgttctgtctggccagaggtgttgggggtctggcttaagcaatggagaag   | 1951 |
| Dh | 3552 | ccttaagacatcaagatctgttctgtctggagagagagctgggggagctggccaaagcaatctgagagag | 3611 |
| Qy | 1952 | tgatgaggggacattctcatcaaaacagttctctgaaagataagtcacagctggcaaaaatctgaa     | 2011 |
| Dh | 3612 | tgatgaggggacattctcatcaaaacagttctctgaaagataagtcacagctggcaaaaatctgaa     | 3671 |
| Qy | 2012 | ccctggaaaacctggagatagaatcgttagagggacccagttcagtcagatcagagccagaga        | 2071 |
| Dh | 3672 | ccctggaaaacctggagatagaatcgttagagggacccagttcagtcagatcagagccagaga        | 3731 |
| Qy | 2072 | aggccccaattctgagagtggtcccccacacccctcttcagacctcttgccgaaatgggttagtg      | 2131 |
| Dh | 3732 | aggccccaattctgagagtggtcccccacacccctcttcagacctcttgccgaaatgggttagtg      | 3791 |
| Qy | 2132 | atcacacacagttcatctctgcagagcaaaatctcacaaagatgttgagcaaaagaggaatgaatttg   | 2191 |
| Dh | 3792 | atcacacacagttcatctctgcagagcaaaatctcacaaagatgttgagcaaaagaggaatgaatttg   | 3851 |
| Qy | 2192 | gtttacagctggaaaaataatacaagagagcgttatggaaaccttaacagtcgagctgcataga       | 2251 |
| Dh | 3852 | gtttacagctggaaaaataatacaagagagcgttatggaaaccttaacagtcgagctgcataga       | 3911 |
| Qy | 2252 | ttgaaactggaaaaaggtccatagtggttttgggccaaagtccttccttgagaaacaagaacgat      | 2311 |
| Dh | 3912 | ttgaaactggaaaaaggtccatagtggttttgggccaaagtccttccttgagaaacaagaacgat      | 3971 |
| Qy | 2312 | ccaagatgagatgtctctcatagctggggatctgacctcaaatggagctgcagagaaaaagatggct    | 2371 |
| Dh | 3972 | ccaagatgagatgtctctcatagctggggatctgacctcaaatggagctgcagagaaaaagatggct    | 4031 |
| Qy | 2372 | gattggcaaatctgcagatagagctctctagagatccaatgtgccaagatttatattggaagaagtc    | 2431 |
| Dh | 4032 | gattggcaaatctgcagatagagctctctagagatccaatgtgccaagatttatattggaagaagtc    | 4091 |
| Qy | 2432 | atcagaatggctctcaatcatcttaaatgtgtccctctcaagctgaagtgaaaaataattcttaca     | 2491 |
| Dh | 4092 | atcagaatggctctcaatcatcttaaatgtgtccctctcaagctgaagtgaaaaataattcttaca     | 4151 |
| Qy | 2492 | gaaataaagaatgcagtgaaatcaagatgggcgtatgtctctggaaaaatggcagtagaaaccttgcc   | 2551 |
| Dh | 4152 | gaaataaagaatgcagtgaaatcaagatgggcgtatgtctctggaaaaatggcagtagaaaccttgcc   | 4211 |
| Qy | 2552 | cttctaactcagaaaatcttcaaaaataagagagacagagccaactgttactactctgagtg         | 2611 |
| Dh | 4212 | cttctaactcagaaaatcttcaaaaataagagagacagagccaactgttactactctgagtg         | 4271 |
| Qy | 2612 | cagctgtgagacctcaagttcaatttaaaaatgtgaaataatctggagcttcccagaagatccag      | 2671 |
| Dh | 4272 | cagctgtgagacctcaagttcaatttaaaaatgtgaaataatctggagcttcccagaagatccag      | 4331 |
| Qy | 2672 | gggggttgggattctgcatactcaagcgaaagaagatacacctcagtggaagtcacataaagaagct    | 2731 |
| Dh | 4332 | gggggttgggattctgcatactcaagcgaaagaagatacacctcagtggaagtcacataaagaagct    | 4391 |
| Qy | 2732 | taaacagagcatvggggtagccagccacggatggagacatcaaaatctcgagatatacgataatctgg   | 2791 |
| Dh | 4392 | taaacagagcatvggggtagccagccacggatggagacatcaaaatctcgagatatacgataatctgg   | 4451 |
| Qy | 2792 | ctgtgataagatgaaatctgtgtgtgtgtgttaacctatggaaagtattatagctcttcgaaga       | 2851 |
| Dh | 4452 | ctgtgataagatgaaatctgtgtgtgtgtgttaacctatggaaagtattatagctcttcgaaga       | 4511 |
| Qy | 2852 | cagcaaaagatgacagttaaaacttaacatccatgcttgagagatccagatcccacagctgtctc      | 2911 |
| Dh | 4512 | cagcaaaagatgacagttaaaacttaacatccatgcttgagagatccagatcccacagctgtctc      | 4571 |
| Qy | 2912 | cttcaagcagctgtgtgcagccagtgagagaaaaaaagaaacagctctccagattctcgatgctcc     | 2971 |

Dh 4572 ctccagcagctcgtgcaagccagtcgagaaanaaagaaacagctcccaagctctcgtatgctcc 4631  
OY 2972 cacaagctcgtctcccaagaaacccggaagctccatccgaataatacagaagcatcatcaacacag 3031  
Dh 4632 cacaagctcgtctcccaagaaacccggaagctccatccgaataatacagaagcatcatcaacacag 4691  
OY 3032 caattctcgtctcgtatctcgtcaaacctcgtcccatatccctcgtcgtcgtcgtcaaacacatcg 3091  
Dh 4692 caattctcgtctcgtatctcgtcaaacctcgtcccatatccctcgtcgtcgtcgtcaaacacatcg 4751  
OY 3092 agattcccaaaagggcgaaacagagctcgtgcctcgtgacatcgtctgggggttcacaaacgtcgc 3151  
Dh 4752 agattcccaaaagggcgaaacagagctcgtgcctcgtgacatcgtctgggggttcacaaacgtcgc 4811  
OY 3152 tgggtgaccttatatacctatgaagtcttaagaaagagagagcagcatgttaaatgtgaagac 3211  
Dh 4812 tgggtgaccttatatacctatgaagtcttaagaaagagagcagcatgttaaatgtgaagac 4871  
OY 3212 tctgggctcgtgagatcagaatcttaagagtgaaatggaatcgtgagaaagggccacacatg 3271  
Dh 4872 tctgggctcgtgagatcagaatcttaagagtgaaatggaatcgtgagaaagggccacacatg 4931  
OY 3272 atggaagcaataaagtctcctgagacagagcgcaagagagtgctgcctgacactctcaacag 3331  
Dh 4932 atggaagcaataaagtctcctgagacagcgcaagagagtgctgcctgacactctcaacag 4991  
OY 3332 atcgagggcccaatacaaaagagagagagtgctgcaacccctcactatctgagctcgaaga 3391  
Dh 4992 atcgagggcccaatacaaaagagagagagtgctgcaacccctcactatctgagctcgaaga 5051  
OY 3392 agccgggaaagagccctagatctaatgctctgctaaagaaacgcatctcgtgagatcttg 3451  
Dh 5052 agccgggaaagagccctagatctaatgctctgctaaagaaacgcatctcgtgagatcttg 5111  
OY 3452 tgcgaagacatgtccaaagagagagaaatctgacagatcccgatcgtgaagagctgatacgaagagcc 3511  
Dh 5112 tgcgaagacatgtccaaagagagagaaatctgacagatcccgatcgtgaagagctgatacgaagagcc 5171  
OY 3512 agataatctcgtgaaatggtgagaaagcgtctgtaaatgctccccaagaaagcgtctgcgctt 3571  
Dh 5172 agataatctcgtgaaatggtgagaaagcgtctgtaaatgctccccaagaaagcgtctgcgctt 5231  
OY 3572 tgcctaaagtgctccctcgtgagccaaatctggaagcttgggaagaaatcaaaactggtccat 3631  
Dh 5232 tgcctaaagtgctccctcgtgagccaaatctggaagcttgggaagaaatcaaaactggtccat 5291  
OY 3632 tccatccagagagagagcgacatccaacacagcgagtgagtgaaagcgagcgtctgctctt 3691  
Dh 5292 tccatccagagagagagcgacatccaacacagcgagtgagtgaaagcgagcgtctgctctt 5351  
OY 3692 tcaatttccatctcgtgatacagatactgagtcacatggaagttagctcaagaaga 3751  
Dh 5352 tcaatttccatctcgtgatacagatactgagtcacatggaagttagctcaagaaga 5411  
OY 3752 atgcatctgcatctgaaatacagagataaagaagatcgaataatgaanaagggccctactg 3811  
Dh 5412 atgcatctgcatctgaaatacagagataaagaagatcgaataatgaanaagggccctactg 5471  
OY 3812 actcaatggaatcagacatcgtgagagagagagcagcccaactcgtgatactgctataat 3871  
Dh 5472 actcaatggaatcagacatcgtgagagagagagcagcccaactcgtgatactgctataat 5531  
OY 3872 ttattgcaatgatacacaacacatggaagtctgacagacagaccnaaaactcgaagtgggg 3931  
Dh 5532 ttattgcaatgatacacaacacatggaagtctgacagacagaccnaaaactcgaagtgggg 5591  
OY 3932 atagagatgtacacatctggtgacatccactgtgggacatgaccccaacgaagtta 3991  
Dh 5592 atagagatgtacacatctggtgacatccactgtgggacatgaccccaacgaagtta 5651  
OY 3992 acccaatgaaatcagatcgtgacatccactggaatcgaagtggtctgctgagagagacgtga 4051  
Dh 5652 acccaatgaaatcagatcgtgacatccactggaatcgaagtggtctgctgagagagacgtga 5711

OY 4052 gttgtgtcacaggtcatcatcaaggagcctggaagttccagatcttcttcaactggtgca 4111  
Dh 5712 gttgtgtcacaggtcatcatcaaggagcctggaagttccagatcttcttcaactggtgca 5771  
OY 4112 cgtcaacacagatatacttaagagatatttaagaaacctccatctgaatgctatacactag 4171  
Dh 5772 cgtcaacacagatatacttaagagatatttaagaaacctccatctgaatgctatacactag 5831  
OY 4172 agcgagagacagatggtcttaagttcagttatagttggaagatatgagccctcaatgag 4231  
Dh 5832 agcgagagacagatggtcttaagttcagttatagttggaagatatgagccctcaatgag 5891  
OY 4232 acttaacattatgataaacaagtgtttcgaagggagcgccctcgaagacgagctc 4291  
Dh 5892 acttaacattatgataaacaagtgtttcgaagggagcgccctcgaagacgagctc 5951  
OY 4292 tgaanaagggcgcatcaagatctgctgcaatggtgagcagagctcgaagagtcacccatg 4351  
Dh 5952 tgaanaagggcgcatcaagatctgctgcaatggtgagcagagctcgaagagtcacccatg 6011  
OY 4352 aagaagctggtgcatactcttaaacggaacaaaggaagcagctgacattgattgtctctt 4411  
Dh 6012 aagaagctggtgcatactcttaaacggaacaaaggaagcagctgacattgattgtctctt 6071  
OY 4412 gaattggtctgcagaatgaaaccaaccccaacccctagctcactcctcactgtaaaagaa 4471  
Dh 6072 gaattggtctgcagaatgaaaccaaccccaacccctagctcactcctcactgtaaaagaa 6131  
OY 4472 tgcactggtcctgacaaatcttatagtctgttcagccgggtcctcaaaactgtaggggg 4531  
Dh 6132 tgcactggtcctgacaaatcttatagtctgttcagccgggtcctcaaaactgtaggggg 6191  
OY 4532 aaataacacttaagttcttcttccatctgaagaaatgcttccctacagacacactaa 4591  
Dh 6192 aaataacacttaagttcttcttccatctgaagaaatgcttccctacagacacactaa 6251  
OY 4592 tcaatttcttcttctcctcgtgacatttggaacttaagaaagaaatattgtgtagtg 4651  
Dh 6252 tcaatttcttcttctcctcgtgacatttggaacttaagaaagaaatattgtgtagtg 6311  
OY 4652 aatctcgttattatttggggagatatcaaatggttttggaagcaatggtgcaagattat 4711  
Dh 6312 aatctcgttattatttggggagatatcaaatggttttggaagcaatggtgcaagattat 6371  
OY 4712 acatgtaagctggttgaataaagaaagataatcttaaaactcaacaaagaaatggtc 4771  
Dh 6372 acatgtaagctggttgaataaagaaagataatcttaaaactcaacaaagaaatggtc 6431  
OY 4772 tcaagtaagttagatgaaataatgaaataataaagaagaatctcggtgagttta 4831  
Dh 6432 tcaagtaagttagatgaaataatgaaataataaagaagaatctcggtgagttta 6491  
OY 4832 aaaaaaagcctcaattggaatctactcctctcccccacaaact 4880  
Dh 6492 aaaaaaagcctcaattggaatctactcctctcccccacaaact 6540

RESULT 3  
AA236453  
ID AA236453 standard; DNA: 7431 BP.  
XX  
AC AA236453:  
XX  
XX  
DT 22-FEB-2000 (first entry)  
XX  
XX  
DE DNA encoding the MMS2 protein.  
XX  
XX Human; MMS2; MNA1; PDZ domain; tumour suppressor; tyrosine phosphatase;  
KW scaffolding protein; cancer; ss.  
XX  
OS Homo sapiens.  
XX



|       |  |  |
|-------|--|--|
| FH    | Key  | Location/Qualifiers  |
| FT    | CDS  | 57..6170   |
| FT    |  | /tag= a  |
| FT    |  | /product= "MMSC2 protein"  |
| XX    |  |  |
| PN    |  | M09958548-A1.  |
| XX    |  |  |
| PD    |  | 18-NOV-1999.   |
| XX    |  |  |
| PF    |  | 07-MAY-1999; 99WO-US09969.   |
| XX    |  |  |
| PR    |  | 08-MAY-1998; 98US-0084740.   |
| XX    |  |  |
| PA    |  | (MYRI-) MYRIAD GENETICS INC.   |
| XX    |  |  |
| P1    |  | Bartel PL, Tavtigian SV;   |
| XX    |  |  |
| DR    |  | WPI: 2000-053077/04.   |
| XX    |  |  |
| PT    |  | P-PSDB: AAY53753.  |
| XX    |  |  |
| PS    |  | Claim 2; Page 84-93; 112pp; English.   |
| XX    |  |  |
| CC    |  | The present sequence encodes the human MMSC2 protein. The MMAC1 protein binds to MMSC2. The MMSC2 protein has 11 post-synaptic density protein, disc-large, zo-1 (PDZ) domains and one or more of these domains interacts specifically with the carboxyl terminal amino acids of MMAC1 (see AAY53754). Specifically, it appears that domain 7, 10 and 13 interact with MMAC1. Since MMSC2 contains 11 PDZ domains and interacts with MMAC1, a known tumour suppressor having a region of homology with protein tyrosine phosphatases, MMSC2 acts as a scaffolding protein in a common biological pathway with MMAC1. It is believed that the interaction between MMAC1 and MMSC2 is required for the tumour suppressor activity of MMAC1. The MMSC2 polypeptides, polynucleotides, fragments and specific or complex specific antibodies may be used for detecting cancer or a predisposition to cancer and screening for agents that may be used to treat MMSC2 and/or MMAC1 related cancer. The polypeptides and polynucleotides may also be used to treat cancer. |
| CC    |  |  |
| CC    |  |  |
| CC    |  |  |
| CC    |  |  |
| CC    |  |  |
| CC    |  |  |
| SQ    |  | Sequence 7431 BP: 2321 A; 1447 C; 1741 G; 1922 T; 0 other:   |
| <hr/> |  |  |
|       | Query Match  | 92.7%; Score 4525.6; DB 21; Length 7431;   |
|       | Best Local Similarity  | 97.4%; Pred. No. 0;  |
|       | Matches 4655; Conservative   | 0; Mismatches 14; Indels 111; Gaps 1   |
| OY    | 212 aggtataatgycataacttacttgagggaatacaccagaatgctgtaatatctaagaag        | 271  |
| Db    | 1858 agtlaaatgycataacttacttgagggaatacaccagaatgctgtaatatctaagaag        | 1917   |
| OY    | 272 aactccctataagaattgaacatgagtgtgtgcctcgcaactctggcacccaccaacctat      | 331  |
| Db    | 1918 aactccctataagaattgaacatgagtgtgtgcctcgcaactctggcacccaccaacctat     | 1977   |
| OY    | 332 cagaatttgatagccttcggaacttatgtgatattgaagctaaacagaanaagcctacgtatgcc  | 391  |
| Db    | 1978 cagaatttgatagccttcggaacttatgtgatattgaagctaaacagaanaagcctacgtatgcc | 2037   |
| OY    | 392 tagtgtgaattcatcgggtgcatcatcagagacaagagatccagtgctgycgatgactgatcgag  | 451  |
| Db    | 2038 tagtgtgaattcatcgggtgcatcatcagagacaagagatccagtgctgycgatgactgatcgag | 2097   |
| OY    | 452 gtccagaatgacagaagaagttccaagacactctggccaatgrygggaagctgcatcacaacaa   | 511  |
| Db    | 2098 gtccagaatgacagaagaagttccaagacactctggccaatgrygggaagctgcatcacaacaa  | 2157   |
| OY    | 512 taagaactggaagaagaaggaacaaagactggttttaaagcattttaagatttcagagatccaa   | 571  |
| Db    | 2158 taagaactggaagaagaaggaacaaagactggttttaaagcattttaagatttcagagatccaa  | 2217   |
| OY    | 572 ttgataccagcaagcactgtagattataattcgcttcttggtgcccctgcgagcatctgtcga    | 631  |

|    |      |  |      |
|----|------|--|------|
| Db | 2218 | ttgtatccaaagacacgvtatataatcttcttctgtgcctcgcgagcatctgtcgaaa             | 2277 |
| Oy | 632  | aggaatgagacgactctctcccggtgacgagatctatgtcttaacagatgttaactctgaaa         | 691  |
| Db | 2278 | aggaatgagacgactctctcccggtgacgagatctatgtcttaacagatgttaactctgaaa         | 2337 |
| Oy | 692  | acagacagctctgagaaagctgtctgaagacactgaaagagacccgtccagagactctgaaa         | 751  |
| Db | 2338 | acagacagctctgagaaagctgtctgaagacactgaaagagacccgtccagagactctgaaa         | 2397 |
| Oy | 752  | taggaagtctgcaaaagccttaaccccttcacacaaagaaaggttaigtcttctgtaagagg         | 811  |
| Db | 2398 | taggaagtctgcaaaagccttcaccccttcacacaaagaaaggttaigtcttctgtaagagg         | 2457 |
| Oy | 812  | atctcttctcttaacccacacacactccctgtctgaagaaacagagctctgtctgaacacccct       | 871  |
| Db | 2458 | atctcttctcttaacccacacacactccctgtctgaagaaacagagctctgtctgaacacccct       | 2517 |
| Oy | 872  | tcaagctgactctgtgctctgtgctgagcaaaatgactgtactatgatagttaatccacat          | 931  |
| Db | 2518 | tcaagctgactctgtgctctgtgctgagcaaaatgactgtactatgatagttaatccacat          | 2577 |
| Oy | 932  | ttgagctctccatctctccctcgaaaatgagacgacatctaacctctacagccttatctat          | 991  |
| Db | 2578 | ttgagctctccatctctccctcgaaaatgagacgacatctaacctctacagccttatctat          | 2637 |
| Oy | 992  | ctctctcatgagagctctctgtgtgtgtgagtcgctgtgaactatgtctctccctctcaatctc       | 1051 |
| Db | 2638 | ctctctcatgagagctctctgtgtgtgtgagtcgctgtgaactatgtctctccctctcaatctc       | 2697 |
| Oy | 1052 | ctctctaaagagatgatatgtaaaaatctctgtgatccagttaactgtactctgatactgtctc       | 1111 |
| Db | 2698 | ctctctaaagagatgatatgtaaaaatctctgtgatccagttaactgtactctgatactgtctc       | 2757 |
| Oy | 1112 | aggaacatataaccacgaatctccctgtgaaagaaagaaatgaagaataacactctgtgtgaca       | 1171 |
| Db | 2758 | aggaacatataaccacgaatctccctgtgaaagaaagaaatgaagaataacactctgtgtgaca       | 2817 |
| Oy | 1172 | taagatctagggccgcgtctctcgtgcttaactataatgaacacacacccctgtgaaaatgtctat     | 1231 |
| Db | 2818 | taagatctagggccgcgtctctcgtgcttaactataatgaacacacacccctgtgaaaatgtctat     | 2877 |
| Oy | 1232 | aacaacaataatgaaatgttaaaaaacaacaataatgtgtgtgaactgaatctcatccaaagtga      | 1291 |
| Db | 2878 | aacaacaataatgaaatgttaaaaaacaacaataatgtgtgtgaactgaatctcatccaaagtga      | 2937 |
| Oy | 1292 | ttataatcaagtgtcagaaactctctctgtgtctacccgaaatctcagctgtgaaaaggtcttgac     | 1351 |
| Db | 2938 | ttataatcaagtgtcagaaactctctctgtgtctacccgaaatctcagctgtgaaaaggtcttgac     | 2997 |
| Oy | 1352 | accgctctgaaacaaagctctccctgtgcctgtgaatgactgtgaagtgtctaatgttaatgtat      | 1411 |
| Db | 2998 | accgctctgaaacaaagctctccctgtgcctgtgaatgactgtgaagtgtgtcatgcttccaaagtgtat | 3057 |
| Oy | 1412 | ctaaagaatctcttgaagaagacataataataagacaagaagacatctctagcttagagaatga       | 1471 |
| Db | 3058 | ctaaagaatctcttgaagaagacataataataataagacaagaagacatctctagcttagagaatga    | 3117 |
| Oy | 1472 | cagttaatgtctataaagaatgtgcttgtgggatagtactctgaaagcatattcatctgaaggtc      | 1531 |
| Db | 3118 | cagttaatgtctataaagaatgtgcttgtgggatagtactctgaaagcatattcatctgaaggtc      | 3177 |
| Oy | 1532 | ccatctagctgagatgagccggaatgtccactgtgagagactgcgaactctgtccatataatgaaggt   | 1591 |
| Db | 3178 | ccatctagctgagatgagccggaatgtccactgtgagagactgcgaactctgtccatataatgaaggt   | 3237 |
| Oy | 1592 | ctaccatcatgtataccaatgtgccacagacagagactatgtctgaagaagacatctctcatctg      | 1651 |
| Db | 3238 | ctaccatcatgtataccaatgtgccacagacagagactatgtctgaagaagacatctctcatctg      | 3297 |
| Oy | 1652 | gcctctgacataaaaattacttatgtgtcctgtcgagaaacatttgtaagagttcaaaaataagct     | 1711 |

Dh 3298 gccctgacataaaaaattactatgtgcttcgacagacatttgaagagatccaaataaagct 3357  
Oy 1712 tgggacaacaactctggaagagtaatgtgcactggaattttctctacacatgagag 1771  
Dh 3358 tgggacaacaactctggaagagtaatgtgcactggaattttctctacacatgagag 3417  
Oy 1772 acaattccagaaattccagagcgagagagagagagagagagagagagagagagagagag 1831  
Dh 3418 acaattccagaaattccagagcgagagagagagagagagagagagagagagagagagag 3477  
Oy 1832 cagcatataagcaatttggataaagccagcggtgtgaaactctgagagagagagagagag 1891  
Dh 3478 cagcatataagcaatttggataaagccagcggtgtgaaactctgagagagagagagag 3537  
Oy 1892 ccttagagatcagcatctgtgtgacagagagatgagagagcggttagagcatgagagag 1951  
Dh 3538 ccttagagatcagcatctgtgtgacagagagatgagagagcggttagagcatgagagag 3597  
Oy 1952 tgaatgagagagcatcttccatacaacaagtctcgtgaagagatgccagctgagcaaaaatgaa 2011  
Dh 3598 tgaatgagagagcatcttccatacaacaagtctcgtgaagagatgccagctgagcaaaaatgaa 3657  
Oy 2012 ccttgaacccgtgagagatgaaatgtaag----- 2039  
Dh 3658 ccttgaacccgtgagagatgaaatgtaag----- 2039  
Oy 2040 ----- 2039  
Dh 3718 atgaaacagctgtgaaagcattcgaagagagagcaacctgtacttattgttaacaga 3777  
Oy 2040 -----ggacccagtcagtcagagatcagagaccagagagagagagagagagagag 2080  
Dh 3778 gcaattataaacaagacaaag 3837  
Oy 2081 tgtgagagtgccccacacacccctccttcagagccttgcagagagagagagagagagagag 2140  
Dh 3838 tgtgagagtgccccacacacccctccttcagagccttgcagagagagagagagagagagag 3897  
Oy 2141 agtcatctgcaagcaaaaattcacaagaagtgtgacaagaagagagagagagagagagagag 2200  
Dh 3898 agtcatctgcaagcaaaaattcacaagaagtgtgacaagaagagagagagagagagagagag 3957  
Oy 2201 ggaataaataatcag 2260  
Dh 3958 ggaataaataatcag 4017  
Oy 2261 agaaaggtcagagtggttggcctaaagctctgtcgtggaacaaagagagagagagagagag 2320  
Dh 4018 agaaaggtcagagtggttggcctaaagctctgtcgtggaacaaagagagagagagagagag 4077  
Oy 2321 ggtctctcatagtgaggatgtgatccaaaatggagctgcagagaaagatgttcgatgtcaca 2380  
Dh 4078 ggtctctcatagtgaggatgtgatccaaaatggagctgcagagaaagatgttcgatgtcaca 4137  
Oy 2381 tttgagagatgagctctcagagatcaatgtgtcagatcttataatggaagagatcagaatg 2440  
Dh 4138 tttgagagatgagctctcagagatcaatgtgtcagatcttataatggaagagatcagaatg 4197  
Oy 2441 cctcatcatcatataaaatgtgtccctctcctcctcctcctcctcctcctcctcctcctcctc 2500  
Dh 4198 cctcatcatcatataaaatgtgtccctctcctcctcctcctcctcctcctcctcctcctcctc 4257  
Oy 2501 atgcagatgaaatcagagtgccgtatgtcctggaatgtcagtaagaaacttgcctcctaact 2560  
Dh 4258 atgcagatgaaatcagagtgccgtatgtcctggaatgtcagtaagaaacttgcctcctaact 4317  
Oy 2561 cagaataatctcaaaataaag 2620  
Dh 4318 cagaataatctcaaaataaag 4377  
Oy 2621 accctgagttatataaaatgtgcaacatctgagagcttcccaagagatcagagagagagag 2680  
Dh 4378 accctgagttatataaaatgtgcaacatctgagagcttcccaagagatcagagagagagag 4437  
Oy 2681 gtaattctacagcgaaagagatacaactcagtggaatcatataaagagagagagagagag 2740  
Dh 4438 gtaattctacagcgaaagagatacaactcagtggaatcatataaagagagagagagagag 4497  
Oy 2741 atgggtagcagcagagatggaag 2800  
Dh 4498 atgggtagcagcagagatggaag 4557  
Oy 2801 atgaaatgt 2860  
Dh 4558 atgaaatgt 4617  
Oy 2861 tgaacataaataccatccatctggaatccagatcccaagagagagagagagagagagagagag 2920  
Dh 4618 tgaacataaataccatccatctggaatccagatcccaagagagagagagagagagagagagag 4677  
Oy 2921 ctggtgcagcagtggaagaaaaaagagagagagagagagagagagagagagagagagagag 2980  
Dh 4678 ctggtgcagcagtggaagaaaaaagagagagagagagagagagagagagagagagagagag 4737  
Oy 2981 gtcocccagagcgag 3040  
Dh 4738 gtcocccagagcgag 4797  
Oy 3041 ctctgagctcgtcaaacctgtgcccataccctgtgctgcagagagagagagagagagagag 3100  
Dh 4798 ctctgagctcgtcaaacctgtgcccataccctgtgctgcagagagagagagagagagagag 4857  
Oy 3101 aagggcgaaacagagctgag 3160  
Dh 4858 aagggcgaaacagagctgag 4917  
Oy 3161 ttattatccatgaagattatgaag 3220  
Dh 4918 ttattatccatgaagattatgaag 4977  
Oy 3221 gagaatcagatctagagagtgaaatgtgaatctgagagagagagagagagagagagagag 3280  
Dh 4978 gagaatcagatctagagagtgaaatgtgaatctgagagagagagagagagagagagagag 5037  
Oy 3281 tcaatgtctcag 3340  
Dh 5038 tcaatgtctcag 5097  
Oy 3341 catacaag 3400  
Dh 5098 catacaag 5157  
Oy 3401 aagggcctagagatgaatgtgtgtaaaagagagagagagagagagagagagagagagag 3460  
Dh 5158 aagggcctagagatgaatgtgtgtaaaagagagagagagagagagagagagagagagag 5217  
Oy 3461 tttgcaaaag 3520  
Dh 5218 tttgcaaaag 5277  
Oy 3521 tttgcaaaag 3580  
Dh 5278 tttgcaaaag 5337  
Oy 3581 gttccctag 3640  
Dh 5338 gttccctag 5397  
Oy 3641 ag 3700  
Dh 5398 ag 5457  
Oy 3701 cactctgtgagatcag 3760  
Dh 5458 cactctgtgagatcag 5517



|    |      |  |      |
|----|------|--|------|
| QY | 3761 | caatcgaatacagaggaattagaacagctcgaataatgaaaaagggccctacacgaactcgaactcgg | 3820 |
| Db | 5518 | catctgaaatacagagggaattagaacagctcgaataatgaaaaagggccctacacgaactcgg     | 5577 |
| QY | 3821 | gaatcagcatcgcgcggaggggtaggcagcccaactctggtagtctgcttatattatcga         | 3880 |
| Db | 5578 | gaatcagcatcgcgcgcggaggggtaggcagcccaactctggtagtctgcttatattatcga       | 5637 |
| QY | 3881 | tgaatgcacccaactcggaggtcttgacagcacaagcccaaaactcgaatttggggaataagtg     | 3940 |
| Db | 5658 | tgaatgcacccaactcggaggtcttgacagcacaagcccaaaactcgaatttggggaataagtg     | 5697 |
| QY | 3941 | tcaacatctgttggcacataccactcgaaggatgatactacacccaagaagttaactactga       | 4000 |
| Db | 5698 | tcacacatctgttggcacataccactcgaaggatgatactacacccaagaagttaactactga      | 5757 |
| QY | 4001 | aaaaatgcatactgcgtccatctggaatctgcaggtctgtctcgtgagagagacgtgagtctga     | 4060 |
| Db | 5758 | aaaaatgcatactgcgtccatctggaatctgcaggtctgtctcgtgagagagacgtgagtctga     | 5817 |
| QY | 4061 | cagatctcaatcgaagagccgcgcgaagttccagatctctctcttcactcggcgcgaactcaaca    | 4120 |
| Db | 5818 | cagatctcaatcgaagagccgcgcgaagttccagatctctctcttcactcggcgcgaactcaaca    | 5877 |
| QY | 4121 | gtatattctcagabtatgttagacccctccctcaatgtaagtctatacactaagacgaagac       | 4180 |
| Db | 5878 | gtatattctcagabtatgttagacccctccctcaatgtaagtctatacactaagacgaagac       | 5937 |
| QY | 4181 | cagatctccttaaggtcttaagatatagtctggagagatagtcgaagccctcatggaagcttaacca  | 4240 |
| Db | 5938 | cagatctccttaaggtcttaagatatagtctggagagatagtcgaagccctcatggaagcttaacca  | 5997 |
| QY | 4241 | tttatctgtaaaaaagctgtcttgcaaaagggagacgcccctgaagaacggaactctgaaaaagg    | 4300 |
| Db | 5998 | tttatctgtaaaaaagctgtcttgcaaaagggagacgcccctgaagaacggaactctgaaaaagg    | 6057 |
| QY | 4301 | ggcatcagatcatctgtctgtcaatcgtgcagagatctagaagaagtcacccatgaaagaagctg    | 4360 |
| Db | 6058 | ggcatcagatcatctgtctgtcaatcgtgcagagatctagaagaagtcacccatgaaagaagctg    | 6117 |
| QY | 4361 | tttgccatcccttaaaacggacaagaagcgacgtctcaactcttgatgtctctctcttgaatctgct  | 4420 |
| Db | 6118 | tttgccatcccttaaaacggacaagaagcgacgtctcaactcttgatgtctctctcttgaatctgct  | 6177 |
| QY | 4421 | ggcgaagattgaacaaacccaacccctcagctccacactctcacttgaagagaatgacatctgt     | 4480 |
| Db | 6178 | ggcgaagattgaacaaacccaacccctcagctccacactctcacttgaagagaatgacatctgt     | 6237 |
| QY | 4481 | ccctgaacaatttcttaagctgtcttcagccgggtctctcaaaaacgtcgaagggggaataaac     | 4540 |
| Db | 6238 | ccctgaacaatttcttaagctgtcttcagccgggtctctcaaaaacgtcgaagggggaataaac     | 6297 |
| QY | 4541 | tttaagttctctttctcatctctagaagaatgctcttccctacgcacaactaacatcatcttctc    | 4600 |
| Db | 6298 | tttaagttctctttctcatctctagaagaatgctcttccctacgcacaactaacatcatcttctc    | 6357 |
| QY | 4601 | tttctctcttcgcatcttgctggaactcttaagaagaaggataattctgttaagtgaaatctgtt    | 4660 |
| Db | 6358 | tttctctcttcgcatcttgctggaactcttaagaagaaggataattctgttaagtgaaatctgtt    | 6417 |
| QY | 4661 | tttatctcgtgagatatactaaatgtttctgtatcacaatggcgaagaattatactacgtcga      | 4720 |
| Db | 6418 | tttatctcgtgagatatactaaatgtttctgtatcacaatggcgaagaattatactacgtcga      | 6477 |
| QY | 4721 | gctcgtctagatataaagaagaataatctctaagcttaacaaagaagaatgcttcagtaagt       | 4780 |
| Db | 6478 | gctcgtctagatataaagaagaataatctctaagcttaacaaagaagaatgcttcagtaaat       | 6533 |
| QY | 4781 | taggaatgaaaaatgaaaaataataaatgaagaagaagaatctcgggagatttaaaaaaatg       | 4840 |
| Db | 6538 | taggaatgaaaaatgaaaaataataaatgaagaagaagaatctcgggagatttaaaaaaatg       | 6597 |
| QY | 4841 | ccctcaatttggcaactctacactctctctccacccccaact                           | 4880 |

| ID                         | Sequence  | Length |
|----------------------------|---|--------|
| DB                         | 6598 cctcaattgcatctacacctctctcccccaccccaact                             | 6637   |
| RESULT                     | 4   |        |
| AAAX29910                  |   |        |
| ID                         | AAAX29910 standard; DNA; 6750 BP.                                       |        |
| XX                         |   |        |
| AC                         | AAAX29910;  |        |
| XX                         |   |        |
| DT                         | 06-JUL-1999 (first entry)   |        |
| XX                         |   |        |
| DE                         | Clone 38-2-1b encoding protein containing PDZ domain.                   |        |
| XX                         |   |        |
| KW                         | PDZ domain; gene expression; human umbilical vascular endothelial cell; |        |
| KW                         | HUVSC; stimulation; tumor necrosis factor; TNF; protein binding;        |        |
| KW                         | cell; proliferation disorder; cancer; ss.                               |        |
| XX                         |   |        |
| OS                         | Homo sapiens.   |        |
| XX                         |   |        |
| PN                         | WO907846-A1.  |        |
| XX                         |   |        |
| PD                         | 18-FEB-1999.  |        |
| XX                         |   |        |
| PF                         | 12-AUG-1998; 98WO-JP03603.  |        |
| XX                         |   |        |
| PR                         | 19-JUN-1998; 98UP-0189944.  |        |
| PR                         | 12-AUG-1997; 97JP-0230356.  |        |
| XX                         |   |        |
| PA                         | (CHUGAI) CHUGAI RES INST MOLECULAR MEDICINE INC.                        |        |
| XX                         |   |        |
| PI                         | Funahashi S, Miyata S;  |        |
| XX                         |   |        |
| DR                         | WPI: 1999-167423/14.  |        |
| P-PSDB:                    | AAV04733.   |        |
| XX                         |   |        |
| PT                         | Protein containing PDZ domain, whose expression is enhanced by TNF      |        |
| PT                         | stimulation - plays an important role in protein/protein                |        |
| PT                         | interactions and is used for screening for proteins for use in          |        |
| PT                         | treatment of cell proliferation disorders such as cancer                |        |
| XX                         |   |        |
| PS                         | Example 7; Page 202-221; 240pp; Japanese.                               |        |
| XX                         |   |        |
| CC                         | This sequence represents clone 38-2-1b which encodes a new protein .    |        |
| CC                         | containing PDZ domains whose expression in human umbilical vascular     |        |
| CC                         | endothelial cells (HUVSC) is enhanced by stimulation with tumor         |        |
| CC                         | necrosis factor (TNF) alpha. The new protein is used to identify        |        |
| CC                         | proteins which bind to it (particularly to the PDZ domains) and the     |        |
| CC                         | genes encoding them, for use in the treatment of cell proliferation     |        |
| CC                         | disorders such as cancer.   |        |
| XX                         |   |        |
| XX                         |   |        |
| SO                         | Sequence 6750 BP; 2054 A; 1385 C; 1634 G; 1677 T; 0 other:              |        |
| Query Match                | 90.9%; Score 4436.2; DB 20; Length 6750;                                |        |
| Best Local Similarity      | 95.5%; Pred. No. 0;   |        |
| Matches 4661; Conservative | 0; Mismatches 8; Indels 210; Gaps                                       |        |
| 1                          |   |        |
| QY                         | 212 aggtaaatgcatcacttacttctggggaataccaccagaatgtgtgtaattcctaaag          | 271    |
| DB                         | 1872 agttaaattgcatcacttacttctggggaataccaccagaatgtgtgtaattcctaaag        | 1931   |
| QY                         | 272 aactgcatcctaaagtgaataatggtgtgtgtcgtgtaactgtgcaccaccaccacat          | 331    |
| DB                         | 1932 aactgcatcctaaagtgaataatggtgtgtgtcgtgtaactgtgcaccaccaccacat         | 1991   |
| QY                         | 332 cgaatctgcatcgtcctggaacttaatgtgataatgtgagcttaacagaanaagcctcagtagatc  | 391    |
| DB                         | 1992 cgaatctgcatcgtcctggaacttaatgtgataatgtgagcttaacagaanaagcctcagtagatc | 2051   |
| QY                         | 392 taggtgagctcatcgggtcatcagaacagaagaatcagctgctgcatgactgatacggt         | 451    |
| DB                         | 2052 taggtgagctcatcgggtcatcagaacagaagaatcagctgctgcatgactgatacggt        | 2111   |

OY 452 gtccagagtacagaagaggttccaagcaccttctgacatctgaggagctggcatcttcagaca 511  
|||||  
Db 2112 gtccagagtacagaagaggttccaagcaccttctgacatctgaggagctggcatcttcagaca 2171  
OY 512 tagagctcggagaaagagagagcaaaaggaacttggttttaagcattttagatatacagatccaa 571  
|||||  
Db 2172 taagctcggagaaagagagagcaaaaggaacttggttttaagcattttagatatacagatccaa 2231  
OY 572 ttgataccagcaagcaactctgataataatctctctgctgctggcgagcttgctgaaa 631  
|||||  
Db 2232 ttgataccagcaagcaactctgataataatctctctgctgctggcgagcttgctgaaa 2291  
OY 632 aggaatgagcaactctctctctgctgagaccgaacatgcttctgaacagatgtaacttgaaa 691  
|||||  
Db 2292 aggaatgagcaactctctctctgctgagaccgaacatgcttctgaacagatgtaacttgaaa 2351  
OY 692 acagcaatctctgagaaagctgtaaaagcactgaaagagacccgttcagggagctgtgaaa 751  
|||||  
Db 2352 acagcaatctctgagaaagctgtaaaagcactgaaagagacccgttcagggagctgtgaaa 2411  
OY 752 taggaatgctgaagccttctaccccttccaacagaaagatgatagttctctgaaggag 811  
|||||  
Db 2412 taggaatgctgaagccttctaccccttccaacagaaagatgatagttctctgaaggag 2471  
OY 812 attccttctctacccaacacacactcctctgaagagagaggtctgctgacaaacccctct 871  
2472 attccttctctacccaacacacactcctctgaagagagaggtctgctgacaaacccctct 2531  
OY 872 tcaagagctgactctgctctgctgagcaaaatgatagtctgactttagatgaatccaat 931  
Db 2532 tcaagagctgactctgctctgctgagcaaaatgatagtctgactttagatgaatccaat 2591  
OY 932 ttgagctcctacactctccctgtaaaatgacagcatctactcactcaagccctctttat 991  
Db 2592 ttgagctcctacactctccctgtaaaatgacagcatctactcactcaagccctctttat 2651  
OY 992 ctctctacatgagcttctctgctgagctgaaactatgcttctccctccatcatctc 1051  
2652 ctctctacatgagcttctctgctgagctgaaactatgcttctccctccatcatctc 2711  
OY 1052 ctctctacatgagcttctctgctgagctgaaactatgcttctccctccatcatctc 1111  
Db 2712 ctctctacatgagcttctctgctgagctgaaactatgcttctccctccatcatctc 2771  
OY 1112 aggaacatataccagcaatctcctgtaaaagagagatgaaacacccctcgttgaca 1171  
Db 2772 aggaacatataccagcaatctcctgtaaaagagagatgaaacacccctcgttgaca 2831  
OY 1172 taagtaatgagctgctctctgcttcttaataaatacacaacctgcaaatgctatg 1231  
Db 2832 taagtaatgagctgctctctgcttcttaataaatacacaacctgcaaatgctatg 2891  
OY 1232 aacaaacatatgaaatgtaaaacaaatagtgtagactgaaatccatcttaccagtgaa 1291  
Db 2892 aacaaacatatgaaatgtaaaacaaatagtgtagactgaaatccatcttaccagtgaa 2951  
OY 1292 ttatatacagctgcaaaactctctctgctgtaaccgaatctcagctggaagagctctgagc 1351  
Db 2952 ttatatacagctgcaaaactctctctgctgtaaccgaatctcagctggaagagctctgagc 3011  
OY 1352 acctgcttgaacagagctcctcctgctgtaagtctgagtgctgcatgcttcaaaatgta 1411  
Db 3012 acctgcttgaacagagctcctcctgctgtaagtctgagtgctgcatgcttcaaaatgta 3071  
OY 1412 cttaagaacatcttctgaaagagatcttaataatagcaaaagcaatctcagcctaggaatga 1471  
Db 3072 cttaagaacatcttctgaaagagatcttaataatagcaaaagcaatctcagcctaggaatga 3131  
OY 1472 cagttagtgcttaataaagatgagcttgaggatgagtgatcgtctgaaagcatatcataggaatg 1531  
Db 3132 cagttagtgcttaataaagatgagcttgaggatgagtgatcgtctgaaagcatatcataggaatg 3191

OY 1532 ccaatagtcagagatgcggagatgccaatgaggagctgcatctgtccatataatgaagat 1591  
|||||  
Db 3192 ccaatagtcagagatgcggagatgccaatgaggagctgcatctgtccatataatgaagat 3251  
OY 1592 ctaccatcagtgtaaaccaatgcccagagagatagtgtagaagaacatctctctatg 1651  
Db 3252 ctaccatcagtgtaaaccaatgcccagagagatagtgtagaagaacatctctctatg 3311  
OY 1652 gccctgacataaaatctactatgctgctgcaagacatctgaaagatccaaataagct 1711  
Db 3312 gccctgacataaaatctactatgctgctgcaagacatctgaaagatccaaataagct 3371  
OY 1712 tgggaacaacaatctcggagagatagtggaactgataatcttctctcaacactgagag 1771  
Db 3372 tgggaacaacaatctcggagagatagtggaactgataatcttctctcaacactgagag 3431  
OY 1772 acattccagaattacccagagagagagagagaggttgaaagaagcaactccaaaca 1831  
Db 3432 acattccagaattacccagagagagagagagaggttgaaagaagcaactccaaaca 3491  
OY 1832 cagcataatgaatctggaatccagccagcggtggaactctggaagagcaacccaat 1891  
Db 3492 cagcataatgaatctggaatccagccagcggtggaactctggaagagcaacccaat 3551  
OY 1892 ccttaagcatcagcatgctgtggtgagcagagatgaggagctcgttaagcaatgagag 1951  
Db 3552 ccttaagcatcagcatgctgtggtgagcagagatgaggagctcgttaagcaatgagag 3611  
OY 1952 tgatgagagggatcttctatcaaaacatgctcggagagatgccaagctggcaaaatgaa 2011  
Db 3612 tgatgagagggatcttctatcaaaacatgctcggagagatgccaagctggcaaaatgaa 3671  
OY 2012 ccttgaaacctggagatagaaatcgtaag----- 2038  
Db 3672 ccttgaaacctggagatagaaatcgtaagagtgatgaaatgaaactgaagatgcaagcc 3731  
OY 2038 ----- 2038  
Db 3732 atgaaacaagctgtgaaagccatctcgaaagcaagcaacccgtatcttatagtacaga 3791  
OY 2039 ----- 2038  
Db 3792 gcaatataaacaagcaagaaatccctctgctcctctgctgacacaaccttaacct 2039  
OY 2039 -----agg 2041  
Db 3852 agtaacaactcagagacactcaacccatctgctgactctctcaaaatcaaaagccagacaag 3911  
OY 2042 caaccagtcagtcagagtcagagccagagagagctccatgtgtcagtgctcccccacc 2101  
Db 3912 caaccagtcagtcagagtcagagccagagagagctccatgtgtcagtgctcccccacc 3971  
OY 2102 ctccctcagccttgcggaaatggtgtagatcaacaacagctcaatctgcgaagcaaatct 2161  
Db 3972 ctccctcagccttgcggaaatggtgtagatcaacaacagctcaatctgcgaagcaaatct 4031  
OY 2162 cacaagaatgtagacaagaagatgagcttggttcaacgctggaaaatatacagaagagct 2221  
Db 4032 cacaagaatgtagacaagaagatgagcttggttcaacgctggaaaatatacagaagagct 4091  
OY 2222 atggaaccccaacagcgagctgcataatgtaactctggaagaggtcatalagtgtctg 2281  
Db 4092 atggaaccccaacagcgagctgcataatgtaactctggaagaggtcatalagtgtctg 4151  
OY 2282 gccataagctctgctggaacaagaagaccgaatccagatgagatgcttctatagttgaggatg 2341  
Db 4152 gccataagctctgctggaacaagaagaccgaatccagatgagatgcttctatagttgaggatg 4211  
OY 2342 atccaaatgagctgcaagaaagatgctgatatgcaaatctgcaagatgcaactctcaaga 2401  
Db 4212 atccaaatgagctgcaagaaagatgctgatatgcaaatctgcaagatgcaactctcaaga 4271  
OY 2402 tcaatgctcagatcttataatgaaagatcatalcagaatgctccatcaatataaatgtg 2461

|    |      |   |      |
|----|------|---|------|
| Dh | 5352 | atcccgatcgaaagctctgatcccaaggagaccagatatattgtgtgaaytggaagaagcttc     | 5411 |
| Qy | 3542 | gtaatgcctcccaagaagacggtctgcgcgctcttcgttaagtgtctccctaaagctaaacct     | 3601 |
| Dh | 5412 | gtaatgcctcccaagaagacggtctgcgcgctcttcgttaagtgtctccctaaagctaaacct     | 5471 |
| Qy | 3602 | tggaagtcttggaagatcaaaagctgtgccttccttcacatccagaggggggcacatccaaacca   | 3661 |
| Dh | 5472 | tcggaagtcttggaagatcaaaagctgtgccttccttcacatccagaggggggcacatccaaacca  | 5531 |
| Qy | 3662 | gccagctctgaaagggcagacgtctctctcttcacatcccttcacacctctctgcgacagatcat   | 3721 |
| Dh | 5532 | gccagctctgaaagggcagacgtctctctcttcacatcccttcacacctctctgcgacagatcat   | 5591 |
| Qy | 3722 | ctgagctcaactcggaaagatgactcaagaagaatgcatctgcatctgaaatacagggtatcaa    | 3781 |
| Dh | 5592 | ctgagctcaactcggaaagatgactcaagaagaatgcatctgcatctgaaatacagggtatcaa    | 5651 |
| Qy | 3782 | gaacacgtctgaataataaaaaggccctcaactcaactctggaaatccgatactgcgtctgaagag  | 3841 |
| Dh | 5652 | gaacacgtctgaataataaaaaggccctcaactcaactctggaaatccgatactgcgtctgaagag  | 5711 |
| Qy | 3842 | taggcagcccaactgtgtgtatgtgcctatatattatgcaaatgtagcaccaccaactgtgagttg  | 3901 |
| Dh | 5712 | taggcagcccaactgtgtgtatgtgcctatatattatgcaaatgtagcaccaccaactgtgagttg  | 5771 |
| Qy | 3902 | cagcagcagaccccaaaaactctagaagcttggggatagagtttgcacactctgttgacatacca   | 3961 |
| Dh | 5772 | cagcagcagaccccaaaaactctagaagcttggggatagagtttgcacactctgttgacatacca   | 5831 |
| Qy | 3962 | ctgagggagatgactcaacaccccaagacggtttgaacctcaatgtaaaaatgcatctggtcccatg | 4021 |
| Dh | 5832 | ctgagggagatgactcaacaccccaagacggtttgaacctcaatgtaaaaatgcatctggtcccatg | 5891 |
| Qy | 4022 | aaatgcagctgtgttcgtcggggagagacgtgtagtgtgttcacagatgcatcaacaggggcctg   | 4081 |
| Dh | 5892 | aaatgcagctgtgttcgtcggggagagacgtgtagtgtgttcacagatgcatcaacaggggcctg   | 5951 |
| Qy | 4082 | caagctccagactctctctctcaactcgggcgtgcagcccaacagatattctcagatgtagtttag  | 4141 |
| Dh | 5952 | caagctccagactctctctctcaactcgggcgtgcagcccaacagatattctcagatgtagtttag  | 6011 |
| Qy | 4142 | gaactctccaaatgtaaaatctcatcaactcagagcagagaccagaaatggtctaaagttcagata  | 4201 |
| Dh | 6012 | gaactctccaaatgtaaaatctcatcaactcagagcagagaccagaaatggtctaaagttcagata  | 6071 |
| Qy | 4202 | tagttgtgggggataatgagacccctcaatgagagacttaccacttatagtttaaaacaggtcttg  | 4261 |
| Dh | 6072 | tagttgtgggggataatgagacccctcaatgagagacttaccacttatagtttaaaacaggtcttg  | 6131 |
| Qy | 4262 | caaaaggagcagcctctgaaagacgagcgtctgaaaaaggggcgatcagatcatctgctgtca     | 4321 |
| Dh | 6132 | caaaaggagcagcctctgaaagacgagcgtctgaaaaaggggcgatcagatcatctgctgtgtca   | 6191 |
| Qy | 4322 | atgggcacaaagctctagaaagggttcaaccataaagaanaagctgttcgcactctttaaagagaaa | 4381 |
| Dh | 6192 | atgggcacaaagctctagaaagggttcaaccataaagaanaagctgttcgcactctttaaagagaaa | 6251 |
| Qy | 4382 | aaggacactgtcaactcttaatgtctctctctgaaatctggcctgcagaaatctgaacaaaccca   | 4441 |
| Dh | 6252 | aaggacactgtcaactcttaatgtctctctctgaaatctggcctgcagaaatctgaacaaaccca   | 6311 |
| Qy | 4442 | ccactagctcaactctctactactgttaaagaatgtcactgtgtcctgcagaaattttatagctgtg | 4501 |
| Dh | 6312 | ccactagctcaactctctactactgttaaagaatgtcactgtgtcctgcagaaattttatagctgtg | 6371 |
| Qy | 4502 | ttcagccggggcttcaaaacagcttaggggggaaataaacctttaagtctctcttccctcaact    | 4561 |
| Dh | 6372 | ttcagccggggcttcaaaacagcttaggggggaaataaacctttaagtctctcttccctcaact    | 6431 |
| Qy | 4562 | agaaatgcttctccttactgcaaacctaacaatacttctctctctctgtgcaattttgtga       | 4621 |
| Dh | 6432 | agaaatgcttctccttactgcaaacctaacaatacttctctctctctgtgcaattttgtga       | 6491 |

|    |      |  |      |
|----|------|--|------|
| QY | 4662 | acttaagaagaagaataatttgcgttagtggaatctcggctttatttcttgcggaatactcaa  | 4681 |
| D6 | 6492 | acttaagaagaagaataatttgcgttagtggaatctcggctttatttcttgcggaatactcaa  | 6551 |
| QY | 4682 | tgcttctgtagcacaatgcygcaagaattattacatgctaaagctggttagtatanaagaaga  | 4741 |
| D6 | 6552 | tgcttctgtagcacaatgcygcaagaattattacatgctaaagctggttagtatanaagaaga  | 6611 |
| QY | 4742 | taattctcaaaagctaaaccaaagaataatggtcttaagttagatgaaataatgaaatat     | 4801 |
| D6 | 6612 | taattctcaaaagctaaaccaaagaataatggtctctcgttagttagatgaaataatgaaatat | 6671 |
| QY | 4802 | aaaaataagaagaanaatctcggggagtttaaaaaaaatgcccaatttgcgcaattact      | 4861 |
| D6 | 6672 | aaaaataagaagaanaatctcggggagtttaaaaaaaatgcccaatttgcgcaattact      | 6731 |
| QY | 4862 | cctctcccccaccccaact  | 4880 |
| D6 | 6732 | cctctcccccaccccaact  | 6750 |

```

RESULT      5
AAA47424
ID   AAA47424 standard; DNA; 3188 BP.

```

AC AAA47424;

DT 20-OCT-2000 (first entry)

DE Sequence encoding human neuron-associated protein

KM Neurotrophin associated protein, NENAP; neurological disorder; epilepsy;  
KM Ischemic cerebrovascular disease; stroke; cerebral neoplasm;  
KM Alzheimer's disease; Pick's disease; Huntington's disease;  
KM dementia; Parkinson's disease; demyelinating disease; meningitis;  
KM prion disease; Kuru; Creutzfeldt-Jakob disease; neurofibrinosis;  
KM Cerebral palsy; muscular dystrophy; central nervous system; CNS;  
KM peripheral nervous system; PMS; myopathy; schizophrenia;  
KM actinic keratosis; arteriosclerosis; atherosclerosis; BMT; BMTs;  
KM cirrhosis; hepatitis; mixed connective tissue disease; MCTD;  
KM myelofibrosis; paroxysmal nocturnal haemoglobinuria; cancer;  
KM autoimmune disease; inflammation; acquired immunodeficiency syndrome;  
KM AIDS; Addison's disease; adult respiratory distress syndrome;  
KM allergy; ankylosing spondylitis; amyloidosis; anaemia; asthma;  
KM Werner syndrome; trauma; human; ds.

OS Homo sapiens

|    |     |                     |
|----|-----|---------------------|
| FH | Key | Location/Qualifiers |
| FT | CDS | 29..2599            |

**/product=** Neuron associated protein

PN WO200034477-A2

PD 15-JUN-2000.

PF 10-DEC-1999;

PR 11-DEC-1998;

PR 09-FEB-1999; 99US-0119365.

XX  
XX  
(TYPE) TAPER BUTT TWO

XX  
XX  
E  
E  
V  
V  
C  
C  
S  
S  
T  
T  
J  
J  
A  
A  
H  
H

PI Lu DAM, Azlmzai Y;

DR WPI; 2000-423423/36

XX

PT New human neuron-associated proteins and polynucleotides encoding them  
PT useful for diagnosis, treatment and prevention of cell proliferative  
PT disorders including cancer, neuronal and neurological disorders  
XX  
PS Claim 9; Page 131-132; 145pp; English.  
XX

PS Claim 9; Page 131-132; 145pp; English.

Human neuron-associated proteins (NEUP) can be used for for treating or preventing a disorder associated with decreased expression or activity of NEUP. Antagonists of NEUP are useful for treating or preventing disorder associated with increased expression or activity of NEUP. NEUP or their fragments or derivatives are useful for treating neurological disorder such as epilepsy, ischemic cerebrovascular disease, stroke, cerebral neoplasms, Alzheimer's disease, Pick's disease, Huntington's disease, dementia and Parkinson's disease. NEUPs are also useful for treating other demyelinating diseases, bacterial and viral meningitis, prion diseases including kuru, Creutzfeldt-Jakob disease, nutritional and metabolic diseases of the nervous system, neurofibromatosis, other developmental disorders of the central nervous system, cerebral palsy, neuroskeletal disorders, autonomic nervous system disorders, cranial nerve disorders, spinal cord diseases, muscular dystrophy and other neuromuscular disorders, peripheral nervous system disorders, inherited, metabolic, endocrine, and toxic myopathies, mental disorders including mood, anxiety and schizophrenic disorders, a cell proliferative disorder such as actinic keratosis, arteriosclerosis, atherosclerosis, bursitis, cirrhosis, hepatitis, mixed connective tissue disease (MCTD), myelofibrosis, paroxysmal nocturnal hemoglobinuria, cancers of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, and an autoimmune/inflammatory disorder such as acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemias, asthma, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal circulation, viral, bacterial, fungal parasitic, protozoal, and helminthic infections, and trauma. This sequence was given the Incyte ID no. 3201/53CBI.

**SQ** Sequence 3188 BP; 990 A; 644 C; 771 G; 783 T; 0 other,

|             |  |
|-------------|--|
| Query match | 57.98; Score 2825; DB 21; Length 3188; |
|-------------|--|

Matches 2831; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 2040 ggcaccagtcagtcagagtcagagccagagaagctccattgtgcagtggtgccccacc 2099

Db 226 ggcacccagtcagtcagagccagagaaggtccatgtgcaagtgtgccccacc 285

QY 2100 ccctcctcagcctttgccgaaatggtagtatcacacacagtcattctgcaagcaaat 2155

Db 286 ccctcctcagccttgcggaatggtagtatcacacacacagtcattctgcagcaaat 345

QY 2160 ctcaagatgtgacaaagagatgatttggltacagctggaatatcagagagcg 2215

Db 346 ctcaagaatgtgacaaagagatgtagttgtgttacagctggaataatcagagagcg 405

2220 ttatggaacctaacagcgagctgcataatgatgaactggagaagtcatagtgtt 2275

Db 406 ttatgaaccctaacagcgagctgcataatgtgaactcgagaaggtcatagtgtt 465

2280 ggccctaagtcctgctgggaacaagaccgatccagatgagtgtctccatagtgggat 2335

Db 466 gggcctaagtcctgtggaacaagaccgatccagatgagtgtctctcatagtgggat 525

QY 2340 tgatccaatggaagtgtcgatgcaatgcaagatgagctctaga 2399

Db 526 tgatccaatggaagctgcgaatgcaaatgcaagatgagcttctaga 585

2400 gaccaatggtcagattttatatggaagaagtcacagaatgcctcatcaatcataatg 2455

Db 586 gatcaatggtcagattttatatgtgaagaagtcatacagaatgcctcatcaatcataaag 645

QY 2460 tggcccttctaagtgaaaaataatttttatcagaataaagatgcagtgaaatcagatggc 2519

|||||  
Db 646 tgcctcttaagtgataataattttatctcgaataaagatgcagtaacagatgyc 705  
Qy 2520 cgtatgctcgtgaataatgcagtagaaccttgcctcttaactcaagaatctcaataa 2579  
Db 706 cgtatgctcgtgaataatgcagtagaaccttgcctcttaactcaagaatctcaataa 765  
Qy 2580 ggaagaagagccaactgtaactctctgtagagctgtaggagcctaagtcttaataa 2639  
Db 766 ggaagaagagccaactgtaactctctgtagagctgtaggagcctaagtcttaataa 825  
Qy 2640 tctgcaacactctgagactccccaagatcagagggtcttgagatctgcatcagcgaga 2699  
Db 826 tctgcaacactctgagactccccaagatcagagggtcttgagatctgcatcagcgaga 885  
Qy 2700 agatcacactcagtgagatcactcaataaagccttaacagagcatgaggtagcagcagga 2759  
Db 886 agatcacactcagtgagatcactcaataaagccttaacagagcatgaggtagcagcagga 945  
Qy 2760 tggagcagctcaagctcgagagatcagatctagctgtagatgtagaattctgtgta 2819  
Db 946 tggagcagctcaagctcgagagatcagatctagctgtagatgtagaattctgtgta 1005  
Qy 2820 cccatctgaaaaagttatctagcctctctgaaagacagagatgacagtaaaacttaacat 2879  
Db 1006 cccatctgaaaaagttatctagcctctctgaaagacagagatgacagtaaaacttaacat 1065  
Qy 2880 ccaatctgagaatccagatccagagctgtctctcaagagctgtgtagcagctgaga 2939  
Db 1066 ccaatctgagaatccagatccagagctgtctctcaagagctgtgtagcagctgaga 1125  
Qy 2940 aaaaaaagaagctccagctctgtagctgtagcagctgtgtagcagctccccaagagagtc 2999  
Db 1126 aaaaaaagaagctccagctctgtagctgtagcagctgtgtagcagctccccaagagagtc 1185  
Qy 3000 cactcgaataatacagagatcactcaacacacagcaatttctgctctgactctgcaactg 3059  
Db 1186 cactcgaataatacagagatcactcaacacacagcaatttctgctctgactctgcaactg 1245  
Qy 3060 ccccatctacccctgctgcgaaaaaacaacatcgagattccaaaaggcgaaacagctgag 3119  
Db 1246 ccccatctacccctgctgcgaaaaaacaacatcgagattccaaaaggcgaaacagctgag 1305  
Qy 3120 cctgagacatctgtggggtctcagacacagctgcggtgctcttatctcaatgaagtta 3179  
Db 1306 cctgagacatctgtggggtctcagacacagctgcggtgctcttatctcaatgaagtta 1365  
Qy 3180 tgaagaagagcagcatctgtaaaagatgaaagactctggctgtagatcaga tcttaagagtc 3239  
Db 1366 tgaagaagagcagcatctgtaaaagatgaaagactctggctgtagatcaga tcttaagagtc 1425  
Qy 3240 gaaatgaaatctgactctgtagaagagccacacatgtagaagaacatctgctgtagacagac 3299  
Db 1426 gaaatgaaatctgactctgtagaagagccacacatgtagaagaacatctgctgtagacagac 1485  
Qy 3300 ggcacagagagctgcgctcgtcagactctacagagatgagggcccttaaaaagaggaagt 3359  
Db 1486 ggcacagagagctgcgctcgtcagactctacagagatgagggcccttaaaaagaggaagt 1445  
Qy 3360 gctgtagacacctcaactatgagctgcagaaagccgggaaaaagcccttaagataat 3419  
Db 1546 gctgtagacacctcaactatgagctgcagaaagccgggaaaaagcccttaagataat 1605  
Qy 3420 tcttggtaaaagaacgatactgagatctgtgtagacatctgcaaaagaggaatctgc 3479  
Db 1606 tcttggtaaaagaacgatactgagatctgtgtagacatctgcaaaagaggaatctgc 1665  
Qy 3480 agatcccgatgagaagctgtccagagagacagagataatctgtagatgtaggagaagcgt 3539  
Db 1666 agatcccgatgagaagctgtccagagagacagagataatctgtagatgtaggagaagcgt 1725  
Qy 3540 tcttaatgctcccaagaagcgtctgtagcttctgtagaattctcccttaagcagagtaac 3599  
|||||

Db 1726 tcttaatgctcccaagaagcgtctgtagcttctgtagaattctcccttaagcagagtaac 1785  
Qy 3600 cttgtagatctggaagaataataaagctggtccatctcaattcaagagagagccatctcaaac 3659  
Db 1786 cttgtagatctggaagaataataaagctggtccatctcaattcaagagagagagccatctcaaac 1845  
Qy 3660 cagcagatctggaagaagcagcctgtctctctcaactctccactctccagatccagatc 3719  
Db 1846 cagcagatctggaagaagcagcctgtctctctcaactctccactctccagatccagatc 1905  
Qy 3720 atctgtagcactggaagaagtagctcaaaagaagaatgcatctgtagatctgaaatacagagat 3779  
Db 1906 atctgtagcactggaagaagtagctcaaaagaagaatgcatctgtagatctgaaatacagagat 1965  
Qy 3780 aagaacagctcgaaatgaaaaagggcctctgtagatcagcgggaatcagagatctgtagag 3839  
Db 1966 aagaacagctcgaaatgaaaaagggcctctgtagatcagcgggaatcagagatctgtagag 2025  
Qy 3840 agtagcagcccaactgtagatctgtagatctcaatctcaatctgtagatctgtagatctgtagat 3899  
Db 2026 agtagcagcccaactgtagatctgtagatctcaatctcaatctgtagatctgtagatctgtagat 2085  
Qy 3900 tgcagcagagcccaaaactcaagatctgtagagatgagatctgtagatctgtagatctgtagat 3959  
Db 2086 tgcagcagagcccaaaactcaagatctgtagagatgagatctgtagatctgtagatctgtagat 2145  
Qy 3960 cactgtagcactgtagatctcaaaagcttaacactcagaaatgcatctgtagatctgtagat 4019  
Db 2146 cactgtagcactgtagatctcaaaagcttaacactcagaaatgcatctgtagatctgtagat 2205  
Qy 4020 tgaatgtagctgtagctgtagagagagctgtagatctgtagatctgtagatctgtagatctgtagat 4079  
Db 2206 tgaatgtagctgtagctgtagagagagctgtagatctgtagatctgtagatctgtagatctgtagat 2265  
Qy 4080 tgcagatctccagctctctctctcaactgtagatctgtagatctgtagatctgtagatctgtagat 4139  
Db 2266 tgcagatctccagctctctctctcaactgtagatctgtagatctgtagatctgtagatctgtagat 2295  
Qy 4140 aggaactccctcaatgtagatctcaactcaactcaactcaactcaactcaactcaactcaactcaact 4199  
Db 2326 aggaactccctcaatgtagatctcaactcaactcaactcaactcaactcaactcaactcaactcaact 2385  
Qy 4200 tatagtgtagagatctgtagcagccctcaatgtagagactcaactcaactcaactcaactcaactcaact 4259  
Db 2386 tatagtgtagagatctgtagcagccctcaatgtagagactcaactcaactcaactcaactcaactcaact 2445  
Qy 4260 tgcagaagagagcagcctctgtagagagagctgtagaagagagagagagagagagagagagagag 4319  
Db 2446 tgcagaagagagcagcctctgtagagagagctgtagaagagagagagagagagagagagagagag 2505  
Qy 4320 caatgtagcagagctctgtag 4379  
Db 2506 caatgtagcagagctctgtag 2565  
Qy 4380 aaaaagcagctcaactctgtagctctctctcaactctcaactctcaactctcaactctcaactctcaact 4439  
Db 2566 aaaaagcagctcaactctgtagctctctctcaactctcaactctcaactctcaactctcaactctcaact 2625  
Qy 4440 aacccctagctcaactctcaactctgtagaagagagagagagagagagagagagagagagagagagag 4499  
Db 2626 aacccctagctcaactctcaactctgtagaagagagagagagagagagagagagagagagagagagag 2685  
Qy 4500 tcttcaagcagctctcaactctgtag 4559  
Db 2686 tcttcaagcagctctcaactctgtag 2745  
Qy 4560 ctgagaatgctctctcaactcagac 4619  
Db 2746 ctgagaatgctctctcaactcagac 2805  
Qy 4620 gaacttaagaagaagaataatctgtag 4679  
Db 2806 gaacttaagaagaagaataatctgtag 2865





QY 3236 aggtgaattgaattgattgagggaagccacacatgatgaagcaatcaatgctcctagac 3295  
 |||||||  
 Db 1157 aggtgaattgaattgattgagggaagccacacatgatgaagcaatcaatgctcctagac 1216  
 QY 3296 agagcgcacagagagagcgccctacacactacacagagatgagcgccataacagaagagag 3355  
 |||||||  
 Db 1217 agagcgcacagagagagcgccctacacactacacagagatgagcgccataacagaagagag 1276  
 QY 3356 aagtgctgaacacccctactatttgatctgacagaagaagccgggaaaaagcgctagat 3415  
 |||||||  
 Db 1277 aagtgctgaacacccctactatttgatctgacagaagaagccgggaaaaagcgctagat 1336  
 QY 3416 gtttttttgtaaaagaagaagatgagttttttgttgaagcattgttgaagaaggagaa 3475  
 |||||||  
 Db 1337 gtttttttgtaaaagaagaagatgagttttttgttgaagcattgttgaagaaggagaa 1396  
 QY 3476 ttgcagatcccgatggaagaagctgacccagggagacagatattttgttgaatggggaag 3535  
 |||||||  
 Db 1397 ttgcagatcccgatggaagaagctgacccagggagacagatattttgttgaatggggaag 1456  
 QY 3536 aagttctgtaattgctccccaagaagcggttgcgcttgcctaagtgctccctagagcaag 3595  
 |||||||  
 Db 1457 aagttctgtaattgctccccaagaagcggttgcgcttgcctaagtgctccctagagcaag 1516  
 QY 3596 taaccttggaattggaagaatacaaaagctgttccattccattcagaaggagagccatctc 3655  
 |||||||  
 Db 1517 taaccttggaattggaagaatacaaaagctgttccattccattcagaaggagagccatctc 1576  
 QY 3656 aaacccagcgagtgagtggaagcgagcgttctcttccattccattccattccattccattcc 3715  
 |||||||  
 Db 1577 aaacccagcgagtgagtggaagcgagcgttctcttccattccattccattccattccattcc 1636  
 QY 3716 gtacatctggaatcctggaagaagtaagctcaaaagaagaatgcatctgcatctgaaatacag 3775  
 |||||||  
 Db 1637 gtacatctggaatcctggaagaagtaagctcaaaagaagaatgcatctgcatctgaaatacag 1696  
 QY 3776 gattaaagaagaagtcgaagaagaagaagcgccctctgactacacgtgggaatgaagatcgctg 3835  
 |||||||  
 Db 1697 gattaaagaagaagtcgaagaagaagaagcgccctctgactacacgtgggaatgaagatcgctg 1756  
 QY 3836 gaaggaatgagcaagccacactggtgcatgctctattttatgtcatgtaccaccaactg 3895  
 |||||||  
 Db 1757 gaaggaatgagcaagccacactggtgcatgctctattttatgtcatgtaccaccaactg 1816  
 QY 3896 gaagtcgaagcagacagacccaaaactcaagagtcggggatgaatgttccacatctgtgca 3955  
 |||||||  
 Db 1817 gaagtcgaagcagacagacccaaaactcaagagtcggggatgaatgttccacatctgtgca 1876  
 QY 3956 catccactgaggcgatgactcac 4015  
 |||||||  
 Db 1877 catccactgaggcgatgactcac 1936  
 QY 4016 ccaatgaagaagcgagtggttgcggaaggaagtcgagtcgcatcagagtcacatcagcag 4075  
 |||||||  
 Db 1937 ccaatgaagaagcgagtggttgcggaaggaagtcgagtcgcatcagagtcacatcagcag 1996  
 QY 4076 agcctgcaagttccagctcttcttccatctgagtcgcatcagcagtcacatcagcag 4135  
 |||||||  
 Db 1997 agcctgcaagttccagctcttcttccatctgagtcgcatcagcagtcacatcagcag 2056  
 QY 4136 atttaagacccctcctcaatgaatgcttatctacatagagcgagacagatggttgcct 4195  
 |||||||  
 Db 2057 atttaagacccctcctcaatgaatgcttatctacatagagcgagacagatggttgcct 2116  
 QY 4196 tcaagtaagtcggaagataagcagccctcatatgagactacacattatgtttaaaacag 4255  
 |||||||  
 Db 2117 tcaagtaagtcggaagataagcagccctcatatgagactacacattatgtttaaaacag 2176  
 QY 4256 tgtttgcaaaaggagcagcctctgaaagcagagcttgaaaaggcgatcagatcag 4315  
 |||||||  
 Db 2177 tgtttgcaaaaggagcagcctctgaaagcagagcttgaaaaggcgatcagatcag 2236

QY 4316 ctgtcaatggcgagagctctagaagagatccacacatgaagaagcggttgcacatccctaac 4375  
 |||||||  
 Db 2237 ctgtcaatggcgagagctctagaagagatccacacatgaagaagcggttgcacatccctaac 2296  
 QY 4376 ggaacaaaaggcaatgtcactttgatgttctctctctgaatctggtcgcagaaattgaacca 4435  
 |||||||  
 Db 2297 ggaacaaaaggcaatgtcactttgatgttctctctctgaatctggtcgcagaaattgaacca 2356  
 QY 4436 acccaacccctagctcactctcactcgttaagaagaatgcatcgttcctgacaattttat 4495  
 |||||||  
 Db 2357 acccaacccctagctcactctcactcgttaagaagaatgcatcgttcctgacaattttat 2416  
 QY 4496 gctgttcaagccgggtcttcaaacctgtagggggaataaacactaagtttcttcc 4555  
 |||||||  
 Db 2417 gctgttcaagccgggtcttcaaacctgtagggggaataaacactaagtttcttcc 2476  
 QY 4556 tcatctagaatgtcttctcactgacacacacacacacacacacacacacacacacacacac 4615  
 |||||||  
 Db 2477 tcatctagaatgtcttctcactgacacacacacacacacacacacacacacacacacacac 2536  
 QY 4616 ttgtgaacttaagaagaagaatattgtgtgagtgatcctgctttattttgttgaagat 4675  
 |||||||  
 Db 2537 ttgtgaacttaagaagaagaatattgtgtgagtgatcctgctttattttgttgaagat 2596  
 QY 4676 atctaattttgttgaatgcatgagcaagaattatcatgctgaagctgttgaat 4735  
 |||||||  
 Db 2597 atctaattttgttgaatgcatgagcaagaattatcatgctgaagctgttgaat 2656  
 QY 4736 gaaagataatcttaagaagctaaccaagaagaatgtctcagtaagttagaatgaatga 4795  
 |||||||  
 Db 2657 gaaagataatcttaagaagctaaccaagaagaatgtctcagtaagttagaatgaatga 2716  
 QY 4796 aaatataaataaagaagaagaatctcggggagtttaaaaaaatgctcctaattggaat 4855  
 |||||||  
 Db 2717 aaatataaataaagaagaagaatctcggggagtttaaaaaaatgctcctaattggaat 2776  
 QY 4856 ctacccctctcccccacccaact 4880  
 |||||||  
 Db 2777 ctacccctctcccccacccaact 2801

RESULT 7  
 ID AAX29911 standard; DNA; 6666 BP.  
 AC AAX29911:  
 XX  
 DT 06-JUL-1999 (first entry)  
 XX  
 DE Clone 38-2-1c encoding protein containing PDZ domain.  
 XX  
 KW PDZ domain; gene expression; human umbilical vascular endothelial cell;  
 KW HUVEC; stimulation; tumour necrosis factor; TNF; protein binding;  
 KW cell; proliferation disorder; cancer; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO907846-A1.  
 XX  
 PD 18-FEB-1999.  
 XX  
 PF 12-AUG-1998; 98MO-JP03603.  
 XX  
 PR 19-JUN-1998; 98JP-0189944.  
 PR 12-AUG-1997; 97JP-0230356.  
 XX  
 PA (CHUG-) CHUGAI RES INST MOLECULAR MEDICINE INC.  
 PI Funahashi S, Miyata S;  
 XX  
 DR WPI; 1999-167423/14.  
 DR P-PSDB; AAY04734.  
 XX





|    |      |   |      |
|----|------|---|------|
| Db | 4804 | agagatgagtttggttacacgctcggaaaaataacagagacgttctatggaaaccttaacag        | 4863 |
| QY | 3083 | caacc-----atcgagattccaaagagcgaaagggcgccctgagacatctgtgg                | 3134 |
| Db | 4864 | cgaagctcatatgatctgaactcggaaaggtcaatgaatgtttcgggacctgaagctctgcgg       | 4923 |
| QY | 3135 | gggttcagacacgctcgtctgggtgaccttaataccatgaagttctatgaagagagcagcc         | 3194 |
| Db | 4924 | gaacaaagaccgattccagagtgagtgcttctcatagtgggatgtgattccaaatggagctgc       | 4983 |
| QY | 315  | atgtaaagatggaagactctgggctggagatcagatctagagtggaatgtgaattgactt          | 3254 |
| Db | 4984 | aggaagaagatgttcgattctgcaaaatctgcagatgtgagctcttcagagatcaaatgtgcagaattt | 5043 |
| QY | 3255 | gaggaagcccccacatgatatgaagcaatcaatgtcccgagacagacgcccagagatgtgcg        | 3314 |
| Db | 5044 | atatggaagagatcatcaagaatgcctcatcaatcaatcaatgaatgtgccctcttaagtgaa       | 5103 |
| QY | 3315 | ccctacacatctacagagatgaagcccatcaaaagagggagadgtgtgtacacccctac           | 3374 |
| Db | 5104 | ataatatttatcaagaaataaagatgcagtgaaatcaagatgtgcgtatgt-----ctcgg         | 5157 |
| QY | 3375 | tattgaagctgcagaagaagccggggaaaaagcctaggatgaattatgtgtgttaaaagaa         | 3434 |
| Db | 5158 | aaatgcagtaagaaccttgctctctcaatcagaanaatcttcaaaataagagacagagacc         | 5217 |
| QY | 3435 | cgatactggaatttttggcagaaactgtgcaaagagaggaattgcgaatcccgatggagag         | 3494 |
| Db | 5218 | aaactgttaactctctgcatgcaacgtctggagcctcaattcatlaaaatgtgcacactc          | 5277 |
| QY | 3495 | actgatacagaggaagccagatatatattgtgtgaatggggagaagacgtctgtcatgtcccca      | 3554 |
| Db | 5278 | ggagcttcccaaggaatcaggggggtgtgtgttatgtctatcagcgaagaatatacactcag        | 5337 |
| QY | 3555 | aggaagcgtgtgcgcttgcctaagtggttccctagcagatgaaccttggaattgggaag           | 3614 |
| Db | 5338 | tggagtcatcataagagcgttaacagagcatctgggtatagacgacagatgtagcagatcaa        | 5397 |
| QY | 3615 | aatacaagctgttcattccattcaatcagagagagagccatctcaaacacagccagtgtaga        | 3674 |
| Db | 5386 | agtcggaatcatgaatactgcgtctgatagtatgaattgtgtgtgttaccatgtgaaa            | 5457 |
| QY | 3675 | aggcagcctgtctctcttcaacttttccactctcttgatccagtaacatctgaatctatgta        | 3734 |
| Db | 5458 | gttatatgacctctcgagacagcaagaatgacagtaaaacttaacatcatcgtctgagaa          | 5517 |
| QY | 3735 | aagtgcctcaagaagaatgcatgtgcatctgaaatacagggatataagaacagtcgaaat          | 3794 |
| Db | 5518 | tccagatcccaagcgtcttcccttcagagcagctgtgtgcagccagtgtagaaaaaagaaacag      | 5577 |
| QY | 3795 | gaaaaagggccctctgaaccacag---ggatacagacatcgcctgggggggtatggagcccc        | 3855 |
| Db | 5578 | ctccagctctctgtgtgtcccaagctctggaatcagcatcgtctgggggtatggagcccc          | 5637 |
| QY | 3852 | actgtgatatgtgcataattatattgcaagatcagcccaactggagattgtgcagcagac          | 3911 |
| Db | 5638 | actgtgtgattgtgcataattatattgcaagatcagcccaactgtgagttgtgcagcagac         | 5697 |
| QY | 3912 | ccaaaactcagattgtgggatataggaattgtcaaccatctgttgacaatacctatgagagcat      | 3971 |
| Db | 5698 | ccaaaactcagattgtgggataaggaattgtcaaccatctgttgacaatacctatgagagcat       | 5757 |
| QY | 3972 | gactacaacccaagcagtttaactactcagaaaatgtcatctgtctccattgtgaaatgcagtc      | 4031 |
| Db | 5758 | gactacaacccaagcagtttaactactcagaaaatgtcatctgtctccattgtgaaatgcagtc      | 5817 |
| QY | 4032 | gtttcctgaggaagacgtgtggtgtgcacagatcaatacaggaagcctgcacagttccag          | 4091 |
| Db | 5818 | gtttcctgaggaagacgtgtggtgtgcacagatcaatacaggaagcctgcacagttccag          | 5877 |
| QY | 4092 | tccttccttcacctgggctgcagctcaaacagtatatttcaggaattgttagaccttcca          | 4151 |

Db 5878 tcttcttctactggtcgcgtacgtcaacagatatttcagagtgatttagagcctctca 5937  
 Oy 4152 atgtatgcttattacacacagagcagagccacatgcttagagcttcagttggtgag 4211  
 Db 5938 atgtatgcttattacacacagagcagagccacatgcttagagcttcagttggtgag 5997  
 Oy 4212 atatgagcagccctcactgagacttaccattatgtttaaacaagcttgcgaaggagc 4271  
 Db 5998 atatgagcagccctcactgagacttaccattatgtttaaacaagcttgcgaaggagc 6057  
 Oy 4272 agcctctgagagcagagcttctgaaaaggcgatcagatcatctgtctgcaatgagag 4331  
 Db 6058 agcctctgagagcagagcttctgaaaaggcgatcagatcatctgtctgcaatgagag 6117  
 Oy 4332 tctagaagagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 4391  
 Db 6118 tctagaagagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 6177  
 Oy 4392 caatttgatgtct 4451  
 Db 6178 caatttgatgtct 6237  
 Oy 4452 accctctctctgtaaaagagatgctgctcctgacatcttctctctctctctctctctct 4511  
 Db 6238 accctctctctgtaaaagagatgctgctcctgacatcttctctctctctctctctctct 6297  
 Oy 4512 tctcaaaactgtaaggag 4571  
 Db 6298 tctcaaaactgtaaggag 6357  
 Oy 4572 tctctactgacaaactaacaatcatttctctctctctctctctctctctctctctctct 4631  
 Db 6358 tctctactgacaaactaacaatcatttctctctctctctctctctctctctctctctct 6417  
 Oy 4632 aaggaatttctgtaagtgaaatcctgcttctctctctctctctctctctctctctctct 4691  
 Db 6418 aaggaatttctgtaagtgaaatcctgcttctctctctctctctctctctctctctctct 6477  
 Oy 4692 tcaacttgagcaagaatttctacatgctgctgctgctgctgctgctgctgctgctgct 4751  
 Db 6478 tcaacttgagcaagaatttctacatgctgctgctgctgctgctgctgctgctgctgct 6537  
 Oy 4752 gctaaccaagaagaatgctgctgctgctgctgctgctgctgctgctgctgctgct 4811  
 Db 6538 gctaaccaagaagaatgctgctgctgctgctgctgctgctgctgctgctgctgct 6597  
 Oy 4812 agaaaaatctcgaggaggtttaaaaaaagcctcaatttggcatttactcctcctccca 4871  
 Db 6598 agaaaaatctcgaggaggtttaaaaaaagcctcaatttggcatttactcctcctccca 6657  
 Oy 4872 ccccaact 4880  
 Db 6658 ccccaact 6666

RESULT 8

AAZ52969 standard; cDNA: 1681 BP.

ID AAZ52969; AAZ52969;  
 AC AAZ52969;  
 XX 14-MAR-2000 (first entry)  
 DE Human prostate tumor cDNA library derived EST fragment #112.  
 KW Pancreas; tumor; EST; expressed sequence tag; human; cytosolic;  
 KW treatment; ds.  
 XX Homo sapiens.  
 OS DE19820190-A1.  
 PN DE19820190-A1.  
 XX PD 04-NOV-1999.

XX 28-APR-1998; 98DE-1020190.  
 XX 28-APR-1998; 98DE-1020190.  
 XX (META-) METAGEN GES GENOMFORSCHUNG MBH.  
 PA Rosenthal A, Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E;  
 DR WPI: 1999-621386/54.  
 DR P-PSDB: AAY74150, AAY74151, AAY74152.  
 PR New human nucleic acid sequences from pancreatic tumors, and related  
 PR proteins.  
 PS Claim 2; Page 274; 502pp; German.  
 CC This invention describes novel polypeptides and their encoding nucleic  
 CC acids derived from human pancreatic tumor tissue which have cytostatic  
 CC activity. The sequences are also useful in producing pharmaceutical  
 CC compositions for treatment of pancreatic tumors. AAZ52858-253014  
 CC represent expressed sequence tag (EST) fragments derived from a human  
 CC pancreatic tumor cDNA library and which encode the proteins represented  
 CC in AAY73814-174252.  
 CC XX  
 SO Sequence 1681 BP; 498 A; 330 C; 445 G; 408 T; 0 other;  
 Query Match 29.4%; Score 1433.4; DB 20; Length 1681;  
 Best Local Similarity 99.1%; Pred. No. 0;  
 Matches 1462; Conservative 0; Mismatches 11; Indels 2; Gaps 2;  
 Oy 2106 ttcacgcttgcggaatgaggtgagatgacacacagatcctgcaagaatcctcaca 2165  
 Db 1 ttcacgcttgcggaatgaggtgagatgacacacagatcctgcaagaatcctcaca 60  
 Oy 2166 agatgtgacaaagaagagtgagttgttacaagctgcaagaatcctcagaagcgttactg 2225  
 Db 61 agatgtgacaaagaagagtgagttgttacaagctgcaagaatcctcagaagcgttactg 120  
 Oy 2226 aacctaacagcgagcgctgcatatgattgaaacttgagaaagagctatagtttggtgct 2285  
 Db 121 aacctaacagcgagcgctgcatatgattgaaacttgagaaagagctatagtttggtgct 180  
 Oy 2286 aagctctgctggaacaaagacgagctcagagatgcttctcacaagtgaggtatgaccc 2345  
 Db 181 aagctctgctggaacaaagacgagctcagagatgcttctcacaagtgaggtatgaccc 240  
 Oy 2346 aatgtgagctgcaagaagaatgctgctgcaaatgctcagaatgagcttctcagaatcaa 2405  
 Db 241 aatgtgagctgcaagaagaatgctgctgcaaatgctcagaatgagcttctcagaatcaa 300  
 Oy 2406 tgtcagaatttatatgtgaaagaatcagacagatgctcctcacaatcattaaatgagcc 2465  
 Db 301 tgtcagaatttatatgtgaaagaatcagacagatgctcctcacaatcattaaatgagcc 360  
 Oy 2466 tctcaaaatgaaaataatttctacagaagaatgctcagaatgaaatgctcagaatgagc 2525  
 Db 361 tctcaaaatgaaaataatttctacagaagaatgctcagaatgaaatgctcagaatgagc 420  
 Oy 2526 tcttgagaatgagtagaacttgcctctcaactcagaaaatcctcaaatgagagac 2585  
 Db 421 tcttgagaatgagtagaacttgcctctcaactcagaaaatcctcaaatgagagac 480  
 Oy 2586 agagccaacttctactcttctgagcagctgctgagcctcagttcatttaaaatgctca 2645  
 Db 481 agagccaacttctactcttctgagcagctgctgagcctcagttcatttaaaatgctca 540  
 Oy 2646 aca-tctgagcttcccaaga-a-tcaggggggttctggtatctgcatcagcgaagaagat 2703  
 Db 541 aca-tctgagcttcccaaga-a-tcaggggggttctggtatctgcatcagcgaagaagat 600  
 Oy 2704 aca-tctgagcttcccaaga-a-tcaggggggttctggtatctgcatcagcgaagaagat 2763

```

|||||
Db 601 aaaccacagcggagctcatctctaaagagcttaacagcagcaggggttagcacaccacgagctgga 660
Oy 2764 cgactcaaaagtcggagatacagatctgctgtatagatgataatctgtctgttaccct 2823
Db 661 cgactcaaaagtcggagatacagatctgctgtatagatgataatctgtctgttaccct 720
Oy 2824 attgaaaagtattattagctctcttgaagacagcaaaagatgacagtaacttaaccatcct 2883
Db 721 attgaaaagttattattagctctcttgaagacagcaaaagatgacagtaacttaaccatcct 780
Oy 2884 gcttgagaacacagatctccagcgtctctctcctcagcagcgtgtcagccagtcgagaanaa 2943
Db 781 gcttgagaacacagatctccagcgtctctctcctcagcagcgtgtcagccagtcgagaanaa 840
Oy 2944 aagaacacagctccagctctctgtatgtgtccacagctgtgtctccacagacaggaatccatc 3003
Db 841 aagaacacagctccagctctctgtatgtgtccacagctgtgtctccacagacaggaatccatc 900
Oy 3004 cgaataacacagatcatcatcaacacacagcaatttttctctctgtatctctcagactctccc 3063
Db 901 cgaataacacagatcatcatcaacacacagcaatttttctctctgtatctctcagactctccc 960
Oy 3064 attatccctgctgctgcaaaacacacacatcagatctccaaagggcgagcagggctgtgctctg 3123
Db 961 attatccctgctgctgcaaaacacacacatcagatctccaaagggcgagcagggctgtgctctg 1020
Oy 3124 agcatctgtgtgggggttcacagacacgctgtctgtgctcttaattatccatgaagttaagaa 3183
Db 1021 agcatctgtgtgggggttcacagacacgctgtctgtgctcttaattatccatgaagttaagaa 1080
Oy 3184 gaagagcgacgacatgttaagaatgtaagaactctgctgctgagatcagatctcttaagggat 3243
Db 1081 gaagagcgacgacatgttaagaatgtaagaactctgctgctgagatcagatctcttaagggat 1140
Oy 3244 ggaattgacttgaggaagacgacacacatgtaagaacatcaatgtctctgagacagacgcca 3303
Db 1141 ggaattgacttgaggaagacgacacacatgtaagaacatcaatgtctctgagacagacgcca 1200
Oy 3304 cagagagctgcgctcgaactctcacaagatgagggcccatcaaaagagaggaagtgtgt 3363
Db 1201 cagagagctgcgctcgaactctcacaagatgagggcccatcaaaagagaggaagtgtgt 1260
Oy 3364 gacacccctactatctgactgtcagaagaagcgaggaagggcctaagatttaattgtt 3423
Db 1261 gacacccctactatctgactgtcagaagaagcgaggaagggcctaagatttaattgtt 1320
Oy 3424 ggttaaaagaaacgatacttgagatctgtctcagacatctgtcaaaagaggaattgcagat 3483
Db 1321 ggttaaaagaaacgatacttgagatctgtctcagacatctgtcaaaagaggaattgcagat 1380
Oy 3484 ccgagatgagaagactgtatccagaggaagacagataattgtgtgaatggggaagcgtcgt 3543
Db 1381 gccgagatgagaagactgtatccagaggaagacagataattgtgtgaatggggaagcgtcgt 1440
Oy 3544 aatgcctcccaagaagcgggttcgagcttgcctaa 3578
Db 1441 aatgcctcccaagaagcgggttcgagcttgcctaa 1475

RESULT 9
AAV69613
ID AAV69613 standard; DNA; 2067 BP.
XX
AC AAV69613;
XX
DT 28-JAN-1999 (first entry)
XX
DE Human secreted protein gene 3 clone HE2FV03.
XX
KW Secreted protein; gene therapy; protein therapy; diagnosis; treatment;
KW central nervous system; CNS; immune system; cancer; trauma; liver;
KW reproductive disorder; congenital malformation; degenerative disease;

```

```

KW inflammatory disease; neoplasia; metabolic disorder; testis; placenta;
KW brain; T cell; spleen; lung; heart; rhabdomyosarcoma; endocrine system;
KW endocrinopathy; endocrine polyglandular syndrome; endocrinoma; sepsis;
KW endocrine ophthalmopathy; osteoclastoma; bacterial infection; bone; ds.
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 116..244
FT FT /*tag= a
FT FT /product= "secreted protein"
FT FT /transl_except= (pos:197..199, aa:Xaa)
FT FT /note= "Xaa-unknown"
FT sig_peptide 116..178
FT FT /*tag= b
FT mat_peptide 179..241
FT FT /*tag= c
PN MO9845712-A2.
XX
PD 15-OCT-1998.
XX
PF 07-APR-1998; 98WO-0506801.
XX
PR 30-MAY-1997; 97US-0048184.
PR 08-APR-1997; 97US-0042726.
PR 08-APR-1997; 97US-0042727.
PR 08-APR-1997; 97US-0042728.
PR 08-APR-1997; 97US-0042754.
PR 08-APR-1997; 97US-004825.
PR 30-MAY-1997; 97US-0048068.
PR 30-MAY-1997; 97US-0048070.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Feng P, NI J, Rosen CA, Ruben SM, Yu G;
XX
DR WPI: 1998-594496/50.
XX
DR P-PSDB: AAW63933.
XX
PT New isolated human genes and secreted polypeptide(s) they encode -
PT useful for the diagnosis and treatment of e.g. cancers, CNS
PT disorders, immune system disorders, inflammatory disease and
PT bacterial infections
XX
PS Claim 4: Pages 102-103; 142pp; English.
XX
XX
XX This sequence represents a nucleic acid molecule designated Gene 3 from
XX the human cDNA clone HE2FV03 (deposited as clone ATCC 97955 and ATCC
XX 209074) which encodes a secreted human protein. This gene is expressed
XX primarily in infant brain, prostate, embryo and to a lesser extent in
XX parathyroid, adrenal gland, tumor and thymus and is useful for
XX diagnosing and treating CNS and immune-system diseases.
XX The invention relates to 20 novel genes and their fragments (AAV69611 to
XX AAV69630) and corresponding secreted proteins (AAW63931 to AAW63950)
XX which are useful for preventing, treating or ameliorating medical
XX conditions e.g. by protein of gene therapy. Also pathological conditions
XX can be diagnosed by determining the amount of the new polypeptides in a
XX sample or by determining the presence of mutations in the
XX polynucleotides. Specific uses are based on which tissues they are most
XX highly expressed in, and include developing products for the diagnosis or
XX treatment of central nervous system (CNS) and immune system diseases,
XX reproductive disorders, cancers, congenital malformations, degenerative
XX diseases, trauma, inflammatory disease, neoplasia, metabolic disorders,
XX diseases in testes, placenta, liver, brain and activated T cells, spleen
XX diseases, lung diseases, heart diseases, rhabdomyosarcoma and disorders
XX of the endocrine system or other endocrinopathies, e.g. endocrine
XX polyglandular syndrome, endocrinoma, and endocrine ophthalmopathy,
XX osteoclastoma and other bone remodelling disorders, bacterial infections
XX and sepsis. The polypeptides are also useful for identifying their
XX binding partners.
XX
XX Sequence 2067 BP; 639 A; 363 C; 407 G; 651 T; 7 other:

```

Query Match 17.5%; Score 854.4; DB 19; Length 2067;  
 Best Local Similarity 96.7%; Pred. No. 3.9e-229;  
 Matches 936; Conservative 2; Mismatches 18; Indels 12; Gaps 6;

```

OY 3925 gttgggga tagatgtgtacacatctgtgtgacatccactgagggagatgactacacccaa 3984
    |||||
DB 308 gttgggga tagatgtgtacacatctgtgtgacatccactgagggagatgactacacccaa 367
OY 3985 gcaatcaaccactgtaaaaaatgcatctgtctccatctgaatgcaagtggtgtgtgagga 4044
    |||||
DB 368 gcaatcaaccactgtaaaaaatgcatctgtctccatctgaatgcaagtggtgtgtgagga 427
OY 4045 gacgtgagtggtgtacaggtgtacatcagagagctgtgcaagttcagagcttcttccact 4104
    |||||
DB 428 gacgtgagtggtgtacaggtgtacatcagagagctgtgcaagttcagagcttcttccact 487
OY 4105 gggctgaagtcacacagatatttcaagatgatttagagactcctcaatgtgaagtctatt 4164
    |||||
DB 488 gggctgaagtcacacagatatttcaagatgatttagagactcctcaatgtgaagtctatt 547
OY 4165 acactaagagcagagacagatgagctgaagctcagatagttg--aggaatggcagcc 4222
    |||||
DB 548 acactaagagcagagacagatgagctgaagctcagatagttg--aggaatggcagcc 607
OY 4223 ctcat-ggagagcttaccattatgtttaaaacagtggtt--gcaaaaggagcagctct- 4278
    |||||
DB 608 ctcatggagagcttaccattatgtttaaaacagtggtt--gcaaaaggagcagctctc 667
OY 4279 gaagacgagctctgtaaaaggggcg----atcagatcatgtgtgtcattggcagagatc 4333
    |||||
DB 668 gaagacgagctctgtaaaaggggcg----atcagatcatgtgtgtcattggcagagatc 727
OY 4334 tagaaggggtacacccatgaagaagctgtgtgcacatccttaaacgggcaaaa-ggcactgtc 4392
    |||||
DB 728 tagaaggggtacacccatgaagaagctgtgtgcacatccttaaacgggcaaaa-ggcactgtc 787
OY 4393 acttgagtggtctctctcttgtaattgtgtccagaaatgaaacacacccctagctca 4452
    |||||
DB 788 acttgagtggtctctctcttgtaattgtgtccagaaatgaaacacacccctagctca 847
OY 4453 cctcctactgttaagaagaatgcatctgtgtcgtgacaaattttatgtgtgttcaacggagt 4512
    |||||
DB 848 cctcctactgttaagaagaatgcatctgtgtcgtgacaaattttatgtgtgttcaacggagt 907
OY 4513 ctcaaaactgttagggggaataaacttaagttcttcttccatctagaagaatgcttt 4572
    |||||
DB 908 ctcaaaactgttagggggaataaacttaagttcttcttccatctagaagaatgcttt 967
OY 4573 ccttactgacaaacttaacatacttttcttctctctgtgcaatttgttgaacttaagaaga 4632
    |||||
DB 968 ccttactgacaaacttaacatacttttcttctctctgtgcaatttgttgaacttaagaaga 1027
OY 4633 aggaatatttctgtgagtgatctcgttttcttcttctgtgtgagatatactaatgtttgtagt 4692
    |||||
DB 1028 aggaatatttctgtgagtgatctcgttttcttcttctgtgtgagatatactaatgtttgtagt 1087
OY 4693 cagatgggcagaagaattatatactgaagctgtgtgagtaataaagaagaataatcttaag 4752
    |||||
DB 1088 cagatgggcagaagaattatatactgaagctgtgtgagtaataaagaagaataatcttaag 1147
OY 4753 cttaaccaaaagaagaatgctcagtaagttagagtgaaataaagaataataataaagaaga 4812
    |||||
DB 1148 cttaaccaaaagaagaatgctcagtaagttagagtgaaataaagaataataataaagaaga 1207
OY 4813 gaaatctcggggaggttaaaaaaaatgctcaattgtggcaatctactcctctccac 4872
    |||||
DB 1208 gaaatctcggggaggttaaaaaaaatgctcaattgtggcaatctactcctctccac 1267
OY 4873 cccaaact 4880
    |||||
DB 1268 cccaaact 1275
  
```

RESULT 10  
 AAX29969  
 ID AAX29969 standard; DNA; 755 BP.  
 XX  
 AC AAX29969;  
 XX  
 DT 06-JUL-1999 (first entry)  
 XX  
 DE PDZ domain-containing protein gene clone FH750.  
 XX  
 KW PDZ domain: gene expression; human umbilical vascular endothelial cell;  
 KW HUVEC; stimulation; tumour necrosis factor; TNF; protein binding; PCR;  
 KW cell; proliferation disorder; cancer; primer; amplification; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN MO9907846-A1.  
 XX  
 PD 18-FEB-1999.  
 XX  
 PE 12-AUG-1998; 98WO-JP03603.  
 XX  
 PR 19-JUN-1998; 98JP-0189944.  
 PR 12-AUG-1997; 97JP-0230356.  
 XX  
 PA (CHUG-) CHUGAI RES INST MOLECULAR MEDICINE INC.  
 XX  
 PI Funahashi S, Miyata S;  
 XX  
 DR WPI: 1999-167423/14.  
 DR P-PSDB; AAY04738.  
 XX  
 PT Protein containing PDZ domain, whose expression is enhanced by TNF  
 PT stimulation - plays an important role in protein/protein  
 PT interactions and is used for screening for proteins for use in  
 PT treatment of cell proliferation disorders such as cancer  
 XX  
 PS Example 7; Page 149-151; 240pp; Japanese.  
 XX  
 CC This sequence represents clone FH750 which encodes a new protein  
 CC containing PDZ domains whose expression in human umbilical vascular  
 CC endothelial cells (HUVEC) is enhanced by stimulation with tumour  
 CC necrosis factor (TNF) alpha. The new protein is used to identify  
 CC proteins which bind to it (particularly to the PDZ domains) and the  
 CC genes encoding them, for use in the treatment of cell proliferation  
 CC disorders such as cancer.  
 XX  
 SQ Sequence 755 BP; 232 A; 142 C; 199 G; 182 T; 0 other;

Query Match 15.4%; Score 753.4; DB 20; Length 755;  
 Best Local Similarity 99.9%; Pred. No. 4.6e-201;  
 Matches 754; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

OY 1310 ttctctctgtctaccgcatcagctggaaggctctgagacactgttgaacagagct 1369
    |||||
DB 1 ttctctctgtctaccgcatcagctggaaggctctgagacactgttgaacagagct 60
OY 1370 cctgtgacctgaatgctgtgagtgctcagcttcaaaatgatactcaagaatctttgaaa 1429
    |||||
DB 61 cctgtgacctgaatgctgtgagtgctcagcttcaaaatgatactcaagaatctttgaaa 120
OY 1430 ggaactataatagcaaaaggcaattctgctcagaagaatgacagtgtaagctataaag 1489
    |||||
DB 121 ggaactataatagcaaaaggcaattctgctcagaagaatgacagtgtaagctataaag 180
OY 1490 atggctggggagatgcttgaagcaattatctcagagtgagcattagtcgaagatggcc 1549
    |||||
DB 181 atggctggggagatgcttgaagcaattatctcagagtgagcattagtcgaagatggcc 240
OY 1550 ggaatgcaatggggagctgctgttccatctaataagaagttcacatcagtgtaacca 1609
  
```

```

Db 241 ggattgcatctggggagcctgcatctgttcacatgaagaagctaccatcagtgtaacca 300
Qy 1610 atgcccagcagcagctatgtgtgaagaacatcttctatgtggccctgcataaata 1669
Db 301 atgcccagcagcagcagctatgtgtgaagaacatcttctatgtggccctgcataaata 360
Qy 1670 ctatgtgctctcagacaccttgggaagagttcaaaataagcttgggacacaaactggaa 1729
Db 361 ctatgtgctctcagacacatcttggaagagttcaaaataagcttgggacacaaactggaa 420
Qy 1730 gagtaatggcacttgatattttcttctacacttgcagagacatctcagaattaccag 1789
Db 421 gagtaatggcacttgatattttcttctacacttgcagagacatctcagaattaccag 480
Qy 1790 agcggaaagagagagaggggttgaagaagcgaacttcaaaacacagatatagcaattgga 1849
Db 481 agcggaaagagagagaggggttgaagaagcgaacttcaaaacacagatatagcaattgga 540
Qy 1850 atcagcccgagcggtgtggaactcttgagagaagaacaaacatccttaggcatcagatcg 1909
Db 541 atcagcccgagcggtgtggaactcttgagagaagaacaaacatccttaggcatcagatcg 600
Qy 1910 ttggttgacgagggatcggggagtcggctaaagcaatggaagtgatgaggggcatcttca 1969
Db 601 ttggttgacgagggatcggggagtcggctaaagcaatggaagtgatgaggggcatcttca 660
Qy 1970 tcaaaacatcttctggaagatagtcacagctgcgcacaaatggaaccttgaaacttgagata 2029
Db 661 tcaaaacatcttctggaagatagtcacagctgcgcacaaatggaaccttgaaacttgagata 720
Qy 2030 gaatcgtagagcaccacagtcacagtcagatcagag 2064
Db 721 gaatcgtagagcaccacagtcacagtcagatcagag 755

```

## RESULT 11

AAx29970  
ID AAX29970 standard; DNA; 865 BP.

AAx29970;

06-JUL-1999 (first entry)

PDZ domain-containing protein gene clone FH850.

PDZ domain; gene expression; human umbilical vascular endothelial cell;

HUVEC; stimulation; tumour necrosis factor; TNF; protein binding; PCR;

cell; proliferation disorder; cancer; primer; amplification; ss.

Homo sapiens.

WO9907846-A1.

18-FEB-1999.

12-AUG-1998; 98WO-JP03603.

19-JUN-1998; 98JP-0189944.

12-AUG-1997; 97JP-0230356.

(CHUGAI) CHUGAI RES INST MOLECULAR MEDICINE INC.

Funahashi S, Miyata S;

WPI; 1999-167423/14.

P-PSDB; AAY04739.

Protein containing PDZ domain, whose expression is enhanced by TNF stimulation - plays an important role in protein/protein interactions and is used for screening for proteins for use in treatment of cell proliferation disorders such as cancer

PS Example 7; Page 152-154; 240pp; Japanese.  
XX This sequence represents clone FH850 which encodes a new protein containing PDZ domains whose expression in human umbilical vascular endothelial cells (HUVEC) is enhanced by stimulation with tumour necrosis factor (TNF) alpha. The new protein is used to identify CC proteins which bind to it (particularly to the PDZ domains) and the CC genes encoding them, for use in the treatment of cell proliferation disorders such as cancer.  
XX Sequence 865 BP; 266 A; 163 C; 230 G; 206 T; 0 other;

Query Match 15.0%; Score 730.4; DB 20; Length 865;  
Best Local Similarity 99.9%; Pred. No. 1,4e-194;  
Matches 731; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

Qy 1310 ttccctctgtgtaccgcatcagctggaaggcctctgagcacctgtctggaacagagct 1369
Db 1 ttccctctgtgtaccgcatcagctggaaggcctctgagcacctgtctggaacagagct 60
Qy 1370 cccctgctgtatctgtgagtggtgtcaatgctcaaaatgtatctaaagaatctttgaaa 1429
Db 61 cccctgctgtatctgtgagtggtgtcaatgctcaaaatgtatctaaagaatctttgaaa 120
Qy 1430 ggaacttaatatagcaaaaggcaatcttagccttaggaatgagctagtgcttaataag 1489
Db 121 ggaacttaatatagcaaaaggcaatcttagccttaggaatgagctagtgcttaataag 180
Qy 1490 atggctctggagatgacgtctcgaagcatattcatatggaagtgccatgacagatgagc 1549
Db 181 atggctctggagatgacgtctcgaagcatattcatatggaagtgccatgacagatgagc 240
Qy 1550 ggattgccattggggagctgacatctgtccatgaatgaagaagctaccatcagtgtaacca 1609
Db 241 ggattgccattggggagctgacatctgtccatgaatgaagaagctaccatcagtgtaacca 300
Qy 1610 atgcccagcagcagcagctatgttgagagaacatctctctcttggccctgcataaata 1669
Db 301 atgcccagcagcagcagctatgttgagagaacatctctctcttggccctgcataaata 360
Qy 1670 ctatgtgctctcagacacatcttgggaagagttcaaaataagcttgggacacaaactggaa 1729
Db 361 ctatgtgctctcagacacatcttgggaagagttcaaaataagcttgggacacaaactggaa 420
Qy 1730 gagtaatggcacttgatattttcttctacacttgcagagacatctcagaattaccag 1789
Db 421 gagtaatggcacttgatattttcttctacacttgcagagacatctcagaattaccag 480
Qy 1790 agcggaaagagagaggggttgaagaagcgaacttcaaaacacagatatagcaattgga 1849
Db 481 agcggaaagagagaggggttgaagaagcgaacttcaaaacacagatatagcaattgga 540
Qy 1850 atcagcccgagcggtgtggaactcttgagagaagaacaaacatccttaggcatcagatcg 1909
Db 541 atcagcccgagcggtgtggaactcttgagagaagaacaaacatccttaggcatcagatcg 600
Qy 1910 ttggttgacgagggatcggggagtcggctaaagcaatggaagtgatgaggggcatcttca 1969
Db 601 ttggttgacgagggatcggggagtcggctaaagcaatggaagtgatgaggggcatcttca 660
Qy 1970 tcaaaacatcttctggaagatagtcacagctgcgcacaaatggaaccttgaaacttgagata 2029
Db 661 tcaaaacatcttctggaagatagtcacagctgcgcacaaatggaaccttgaaacttgagata 720
Qy 2030 gaatcgtagag 2041
Db 721 gaatcgtagag 732

```

## RESULT 12

AAx29971  
ID AAX29971 standard; DNA; 959 BP.

```

XX AC AAX29971;
XX DF 06-JUL-1999 (first entry)
XX DE PDZ domain-containing protein gene clone FH950.
XX KW PDZ domain: gene expression; human umbilical vascular endothelial cell;
XX KW HUVEC; stimulation; tumour necrosis factor; TNF; protein binding; PCR;
XX KW cell; proliferation disorder; cancer; primer; amplification; ss.
XX OS Homo sapiens.
XX PN MO9907846-A1.
XX PD 18-FEB-1999.
XX PF 12-AUG-1998; 98WO-JP03603.
XX PR 19-JUN-1998; 98JP-0189944.
XX PR 12-AUG-1997; 97JP-0230356.
XX PA (CHUG-) CHUGAI RES INST MOLECULAR MEDICINE INC.
XX PI Funahashi S, Miyata S;
XX DR WPI: 1999-167423/14.
XX DR P-PSDB: AAY04740.
XX PT Protein containing PDZ domain, whose expression is enhanced by TNF
XX PT stimulation - plays an important role in protein/protein
XX PT interactions and is used for screening for proteins for use in
XX PT treatment of cell proliferation disorders such as cancer.
XX PS Example 7; Page 155-158; 240pp; Japanese.
XX CC This sequence represents clone FH950 which encodes a new protein
XX CC containing PDZ domains whose expression in human umbilical vascular
XX CC endothelial cells (HUVEC) is enhanced by stimulation with tumour
XX CC necrosis factor (TNF) alpha. The new protein is used to identify
XX CC proteins which bind to it (particularly to the PDZ domains) and the
XX CC genes encoding them, for use in the treatment of cell proliferation
XX CC disorders such as cancer.
XX SQ Sequence 959 BP; 295 A; 198 C; 239 G; 227 T; 0 other.

Query Match 15.0%; Score 730.4; DB 20; Length 959;
Best Local Similarity 99.9%; Pred. No. 1.5e-194;
Matches 731; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1310 ttctctctgtgtaaccggtatcagctggaagggctctgagacctgtgtaacagagt 1369
DB 1 ttctctctgtgtaaccggtatcagctggaagggctctgagacctgtgtaacagagt 60
OY 1370 cccctggtctgaagctgagtggtgtcattgcttcaaaatgataaagaarcttttgaa 1429
DB 61 cccctggtctgaagctgagtggtgtcattgcttcaaaatgataaagaarcttttgaa 120
OY 1430 ggaactataataatagcaaaaggaactctcagctcagaagatcaggttgaataaag 1489
DB 121 ggaactataataatagcaaaaggaactctcagctcagaagatcaggttgaataaag 180
OY 1490 atggcttggaatgatcgttcgaagcattatcattgagggagcattatgtaagtgcc 1549
DB 181 atggcttggaatgatcgttcgaagcattatcattgagggagcattatgtaagtgcc 240
OY 1550 ggaattgcatgggagctgactctgttcattaaatgaagagctacacacagtgtaacca 1609
DB 241 ggaattgcatgggagctgactctgttcattaaatgaagagctacacacagtgtaacca 300
OY 1610 atggccagggcagcagctatgttggaagacattcttcattggcccgactaaataa 1669
|||||

```

```

DB 301 atgccagcgacagctatgttggaagacattctctcatttgcgccttgacataaata 360
OY 1670 ctatgtgctcgcgcgaacatttgaagagttcaaaataagcttgggcaaaacttgaa 1729
DB 361 ctatgtgctcgcgcgaacatttgaagagttcaaaataagcttgggcaaaacttgaa 420
OY 1730 gagaatgacacatgatattttcttcatacacttgcagagacattccagaattacag 1789
DB 421 gagaatgacacatgatattttcttcatacacttgcagagacattccagaattacag 480
OY 1790 agcgagaagagggaggtgtaagaagcgaactcaaaacacagcatatagcaattgga 1849
DB 481 agcgagaagagggaggtgtaagaagcgaactcaaaacacagcatatagcaattgga 540
OY 1850 atcagccagcggtgtggaactctggaagagacccaacacttctgagcaccattg 1909
DB 541 atcagccagcggtgtggaactctggaagagacccaacacttctgagcaccattg 600
OY 1910 ttggtgagcagaggtggtggtcgtcgaagcaatggaagtgtatgagggcatttca 1969
DB 601 ttggtgagcagaggtggtggtcgtcgaagcaatggaagtgtatgagggcatttca 660
OY 1970 tcaaacatgtctggaagatagtcacagctggcaaaatggaacttgaaccttgagata 2029
DB 661 tcaaacatgtctggaagatagtcacagctggcaaaatggaacttgaaccttgagata 720
OY 2030 gaatcgtagag 2041
DB 721 gaatcgtagag 732

RESULT 13
AAX29967
ID AAX29967 standard; DNA: 1776 BP.
XX AC AAX29967;
XX DT 06-JUL-1999 (first entry)
XX DE PDZ domain-containing protein gene clone D-2.
XX KW PDZ domain: gene expression; human umbilical vascular endothelial cell;
XX KW HUVEC; stimulation; tumour necrosis factor; TNF; protein binding; PCR;
XX KW cell; proliferation disorder; cancer; primer; amplification; ss.
XX OS Homo sapiens.
XX PN MO9907846-A1.
XX PD 18-FEB-1999.
XX PF 12-AUG-1998; 98WO-JP03603.
XX PR 19-JUN-1998; 98JP-0189944.
XX PR 12-AUG-1997; 97JP-0230356.
XX PA (CHUG-) CHUGAI RES INST MOLECULAR MEDICINE INC.
XX PI Funahashi S, Miyata S;
XX DR WPI: 1999-167423/14.
XX DR P-PSDB: AAY04746.
XX PT Protein containing PDZ domain, whose expression is enhanced by TNF
XX PT stimulation - plays an important role in protein/protein
XX PT interactions and is used for screening for proteins for use in
XX PT treatment of cell proliferation disorders such as cancer.
XX PS Example 7; Page 140-146; 240pp; Japanese.
XX CC This sequence represents clone D-2 which encodes a new protein containing
XX CC PDZ domains whose expression in human umbilical vascular endothelial
XX CC cells (HUVEC) is enhanced by stimulation with tumour necrosis factor

```





```

OY 3466 aagaggaatgctcagatcccgatggaagactgtccagggaagaccagataattgtg 3525
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 4804 aagcgagagcgagccgagctgagatggagatgttcaggagagatcagatctctgtg 4863
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 3526 aatggggaagagctgtctaaagctcccaagaagcglttgcgcttctgtaagtgtcc 3585
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 4864 aatggggaagagctgtctaaagctcccaagaagcglttgcgcttctgtaagtgtca 4923
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 3586 ctgaagcagatgaactgtgaaggttggaaagaaacaaagctgtccatccatcagagag 3645
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 4924 cagggaacttggcagcagatgagatggaaagactccgagctgtctcgtgactccgcaag 4983
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 3646 aggcacatcccaaacccagcagtgagt-----gaagcagccgtctcttccatttc 3700
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 4984 acgacatccacagaaacagtcagaggtacagaggttcacacagcagctgtcaccctcc 5043
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 3701 cactctctgataccagtaactctgagtcacgtggaagtaagtcaagaagaatgcatgtg 3760
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 5044 ctgcctctgataccagtaactctgagtcacgtggaagtaagtcaagaagaatgcatgtc 5103
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 3761 catctgaata-----caggatlaagaaacagtcgaatacgaataaagagccct 3807
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 5104 tccagagaanaatctcagggcacagatagtaaaccaagagactgttgatataaagagagctc 5163
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 3808 actgactcaactggaatcagatcagatcgtctgagagtaggacagccactgtgtatgtcct 3867
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 5164 agtgaatccctctgagatcagatctatgtctggaagagagagcttcttagagatataccc 5223
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 3868 atattatctgcaatgacatgaccccaactgagctgacagacagcccaaaaactagaagct 3927
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 5224 gtaattcttgcagatcagatcagagctgagcagagctgtgcagacagacagaaacttaagct 5283
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 3928 ggggagatgagatgttcacacatctgtgacacatccactgagggatgacatcacaccagca 3987
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 5284 ggaagatcagatgtcagacatcaatcaacgagcaacacttggatgtgctgtccacgagatgtg 5343
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 3988 gttactactcgaanaaatgtacatcgtgctcattgaatgcaggggtgtgtcgaggaggagac 4047
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 5344 gttactactcgaanaaatgtacatcgtgctcattgaatgcaggggtgtgtcgaggaggagac 5403
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 4048 gtgagctgtgcaatgacatcagatcagagagcctgcaagcttccatcttcttcaactggg 4107
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 5404 ataaagcgcaatagacagctcgtgaaaa-----catgtctaaagcttaccacttggct 5457
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 4108 ctgacgtcaaacagtatattcagagatgattaggaacctcctcaatgttaagtctatata 4167
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 5458 tcgcccactgtgaaacacacatccagaaacagaaacacactccaactaagattattact 5517
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 4168 ctgaagcagagacacagatgtgttaggtctcagatagtttgagagatatgcaagccctcat 4227
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 5518 ctgagagaagagctctgaaagcttgggtttagtattgttaggggttattggaagctcccat 5577
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 4228 ggaagacacacatcttattgtaaacacagctgttgcagaagagagagagcctcgaagacgga 4287
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 5578 ggaagacacacatcttattgtaaacacagctgttgcagaagagagagagcctcgaagacgga 5637
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 4288 cgtctgaaaggggagctcagatcagatctgtctgacagagctcagaagagagctcagac 4347
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 5638 cgtatataaagagggagctcagatctttagctgttaatgtgcagagcccggaaggtgtact 5697
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 4348 catgaagaagagctgtctcagatccctaaacagaaagagcagctgtcacttgaatgtctc 4407
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 5698 catgaagaagagctgtctcagatccctaaacagaaagagcagctgtcacttgaatgtctc 5757
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 4408 tcttgaatttgctgcagaa 4426
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 5758 tcatgagctcagggccgaa 5776
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 15  
 AAC27886  
 ID AAC27886 standard; cDNA; 229 BP.  
 XX

```

AC AAC27886;
XX
DE 06-OCT-2000 (first entry)
XX
DE Human secreted protein 5' EST, SEQ ID NO: 31961.
XX
KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KM gene therapy; chromosome mapping; ss.
XX
OS Homo sapiens.
XX
PN EPI033401-A2.
XX
PD 06-SEP-2000.
XX
PF 21-FEB-2000; 2000EP-0200610.
XX
PR 26-FEB-1999; 9905-0122487.
XX
PA (GENST) GENSET.
PI Dumas Milne Edwards J, Duclert A, Giordano J;
XX
DR WPI; 2000-500381/45.
XX
PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX
PS Claim 1; SEQ ID 31961; 71bp + CD-ROM; English.
XX
CC The present sequence is one of a large number of 5' ESTs derived from
CC mRNAs encoding secreted proteins. No ORF has yet been conclusively
CC identified within the present sequence. The 5' ESTs were prepared from
CC total human RNAs or poly(A) RNAs derived from 30 different tissues. EST
CC sequences usually correspond mainly to the 3' untranslated region (UTR)
CC of the mRNA because they are often obtained from oligo-dT primed cDNA
CC libraries. Such ESTs are not well suited for isolating cDNA sequences
CC derived from the 5' ends of mRNAs and even in those cases where longer
CC cDNA sequences have been obtained, the full 5' UTR is rarely included.
CC 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be
CC used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used
CC in diagnostic, forensic, gene therapy and chromosome mapping procedures.
CC They are used to obtain upstream regulatory sequences and to design
CC expression and secretion vectors.
XX
SQ Sequence 229 BP; 99 A; 41 C; 27 G; 58 T; 4 other;

```

Query Match 2.58; Score 124; DB 21; Length 229;  
 Best Local Similarity 96.9%; Pred. No. 9e-25;  
 Matches 124; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

```

OY 4752 gctaaccaagaagaatgtctcagtagtagatggaataatgaaataataaaga 4811
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1 gctaaccaagaagaatgtctcagtagtagatggaataatgaaataataaaga 60
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 4812 agaataatcctgggagagtttaaaaaatgctcattgtgcaatcctcctcccca 4871
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 61 agaataatcctgggagagtttaaaaaatgctcattgtgcaatcctcctcccca 120
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 4872 cccccaac 4879
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 121 mcccccaam 128
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

Search completed: July 12, 2001, 21:53:11  
 Job time: 10454 sec



Fri Jul 13 15:00:16 2001

us-09-502-698-3.rng

W

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

## OM nucleic - nucleic search, using sw model

Run on: July 12, 2001, 18:56:47 ; Search time 112.56 Seconds  
(without alignments)  
8031.898 Million cell updates/sec

Title: US-09-502-698-3

Perfect score: 4880  
Sequence: 1 cccgggccccgggagacagtg.....tctctcccccacccaact 4880

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 317530 seqs, 92630169 residues

Total number of hits satisfying chosen parameters: 635060

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_NA:\*

1: /cgn2\_6/ptodata/2/ina/5A.COMB.seq:\*  
2: /cgn2\_6/ptodata/2/ina/6A.COMB.seq:\*  
3: /cgn2\_6/ptodata/2/ina/6A.COMB.seq:\*  
4: /cgn2\_6/ptodata/2/ina/6B.COMB.seq:\*  
5: /cgn2\_6/ptodata/2/ina/PCITUS.COMB.seq:\*  
6: /cgn2\_6/ptodata/2/ina/Backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description       |
|------------|-------|-------------|--------|----|-------------------|
| 1          | 66    | 1.4         | 8043   | 5  | PCT-US94-09943-1  |
| 2          | 66    | 1.4         | 8119   | 4  | US-09-290-640-45  |
| 3          | 65.6  | 1.3         | 7218   | 1  | US-08-232-463-14  |
| 4          | 64.4  | 1.3         | 8040   | 1  | US-08-596-291-1   |
| 5          | 64.4  | 1.3         | 8040   | 3  | US-09-100-804-1   |
| 6          | 63.6  | 1.3         | 712    | 1  | US-08-410-804-4   |
| 7          | 63.6  | 1.3         | 712    | 1  | US-08-259-514-4   |
| 8          | 63.6  | 1.3         | 712    | 2  | US-08-858-311-4   |
| 9          | 63.6  | 1.3         | 1830   | 1  | US-08-410-804-2   |
| 10         | 63.6  | 1.3         | 1830   | 1  | US-08-259-514-2   |
| 11         | 63.6  | 1.3         | 1830   | 2  | US-08-858-311-2   |
| 12         | 56    | 1.1         | 1396   | 2  | US-09-151-611-2   |
| 13         | 52.8  | 1.1         | 1042   | 4  | US-09-276-531-77  |
| 14         | 49.4  | 1.0         | 2150   | 1  | US-08-580-680-2   |
| 15         | 49.4  | 1.0         | 2150   | 1  | US-08-480-156A-2  |
| 16         | 49.4  | 1.0         | 2150   | 1  | US-08-354-961-2   |
| 17         | 49.4  | 1.0         | 2150   | 2  | US-08-581-094-2   |
| 18         | 49.4  | 1.0         | 2150   | 2  | US-08-580-665-2   |
| 19         | 49.4  | 1.0         | 2150   | 2  | US-08-581-142-2   |
| 20         | 49.4  | 1.0         | 2150   | 2  | US-08-581-103-2   |
| 21         | 49.4  | 1.0         | 2150   | 5  | PCT-US94-05442A-2 |
| 22         | 49.4  | 1.0         | 2151   | 2  | US-08-944-449-3   |
| 23         | 47.8  | 1.0         | 1178   | 4  | US-09-091-405-1   |
| 24         | 46.6  | 1.0         | 2058   | 4  | US-08-749-391-1   |
| 25         | 46.6  | 1.0         | 2058   | 4  | US-09-390-200-1   |
| 26         | 42.6  | 0.9         | 393    | 2  | US-08-944-449-1   |
| 27         | 42.4  | 0.9         | 5852   | 1  | US-07-867-106-2   |

|    |      |     |       |   |                   |                   |
|----|------|-----|-------|---|-------------------|-------------------|
| 28 | 42.2 | 0.9 | 2389  | 1 | US-08-123-161A-13 | Sequence 13, Appl |
| 29 | 42.2 | 0.9 | 2389  | 1 | US-08-483-278-13  | Sequence 13, Appl |
| 30 | 42   | 0.9 | 2447  | 2 | US-09-014-969-14  | Sequence 14, Appl |
| 31 | 40.4 | 0.8 | 7218  | 1 | US-08-232-463-14  | Sequence 1, Appl  |
| 32 | 40.2 | 0.8 | 4527  | 3 | US-09-045-632-1   | Sequence 1, Appl  |
| 33 | 39.8 | 0.8 | 19124 | 2 | US-08-487-826B-13 | Sequence 13, Appl |
| 34 | 39.4 | 0.8 | 1474  | 4 | US-08-821-994-64  | Sequence 64, Appl |
| 35 | 38.6 | 0.8 | 2040  | 3 | US-08-997-445D-1  | Sequence 1, Appl  |
| 36 | 38.4 | 0.8 | 51952 | 4 | US-08-947-823-1   | Sequence 1, Appl  |
| 37 | 37.6 | 0.8 | 642   | 3 | US-08-946-026-33  | Sequence 33, Appl |
| 38 | 37.6 | 0.8 | 2026  | 2 | US-08-993-228-3   | Sequence 3, Appl  |
| 39 | 37.4 | 0.8 | 3705  | 5 | PCT-US96-03940-7  | Sequence 7, Appl  |
| 40 | 37.4 | 0.8 | 5648  | 5 | PCT-US96-03940-1  | Sequence 1, Appl  |
| 41 | 37.2 | 0.8 | 602   | 1 | US-08-764-100-8   | Sequence 8, Appl  |
| 42 | 37.2 | 0.8 | 642   | 1 | US-08-764-100-13  | Sequence 13, Appl |
| 43 | 37.2 | 0.8 | 643   | 1 | US-08-764-100-7   | Sequence 7, Appl  |
| 44 | 37.2 | 0.8 | 2993  | 1 | US-08-764-100-2   | Sequence 2, Appl  |
| 45 | 37.2 | 0.8 | 2993  | 1 | US-08-764-100-10  | Sequence 10, Appl |

## ALIGNMENTS

RESULT 1  
PCT-US94-09943-1  
Sequence 1, Application PC/TUS9409943  
GENERAL INFORMATION:  
APPLICANT:  
APPLICANT:  
APPLICANT:  
APPLICANT:  
TITLE OF INVENTION: PRIMARY STRUCTURE AND FUNCTIONAL  
TITLE OF INVENTION: EXPRESSION OF NUCLEOTIDE SEQUENCES FOR NOVEL PROTEIN  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
STREET: 600 ATLANTIC AVENUE  
CITY: BOSTON  
STATE: MASSACHUSETTS  
COUNTRY: USA  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US94/09943  
FILING DATE: 01-SEP-1994  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/115,573  
FILING DATE: 01-SEP-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: TWOMEY, MICHAEL J.  
REGISTRATION NUMBER: P-38,349  
REFERENCE/DOCKET NUMBER: L0461/7000W0  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617/720-3500  
TELEFAX: 617/720-2441  
TELEX: 92-1742 EZEKIL  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 8043 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA to mRNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: HOMO SAPIENS  
FEATURE:

| Query Match           | 1.4%   | Score 66          | DB 4     | Length 8119 |
|-----------------------|--|-------------------|----------|-------------|
| Best Local Similarity | 58.8%  | Pred. No. 7.5e-09 |          |             |
| Matches 114           | Conservative 0   | Mismatches 80     | Indels 0 | Gaps 0      |
| QY 4187               | gcttagcgttcagatagttgagagatattgcaagccctcatgtagagactaccattatg      | 4246              |          |             |
| Db 4193               | gcttgaggatagtgctcagcggaggtggaatacagagtgctcagacatggtgcattatg      | 4252              |          |             |
| QY 4247               | ttaaaaacagtgcttgcgaaggagcagcccttgaagaacggaagcttgaagaaggcgatc     | 4306              |          |             |
| Db 4253               | tgaagaagctgattatcccccagggagcagcagagctcgtgagtagaattatcacaaggctgac | 4312              |          |             |
| QY 4307               | agatcagctgcttcacatggtgcagagctcagaagaaggtcacccatgaagaagctgtgtcca  | 4366              |          |             |
| Db 4313               | gcgcctcagctgctcaatggaattgcttcagaagaaggccaccatgaagaagctgtggaaa    | 4372              |          |             |

|    |      |  |      |
|----|------|--|------|
| QY | 1685 | aacattcggagaagcttcaataaagcttggacaacaactcgtgaagatgaatgacattg  | 1744 |
|    |      |  |      |
| Db | 1456 | AAGGATAGAGAATTTGGTACRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR     | 1394 |
| QY | 1745 | atatcttctcatalacactcgtgcagagacattcagaattaccagaacgagaagagga   | 1804 |
|    |      |  |      |
| Db | 1396 | RR | 1337 |
| QY | 1805 | aaggctgaagaagcgcaacttcaaacacagcatatagcaattcgaatcagcccaagcgag | 1866 |
|    |      |  |      |
| Db | 1336 | RR | 1277 |
| QY | 1865 | tggaaactcggagaagaccacaagcaatccttggcatcagcattgtgtgtgacgaagga  | 1924 |
|    |      |  |      |

```

; LENGTH: 8040 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; HYPOTHETICAL: NO

```





|             |            |              |           |            |        |        |   |      |   |
|-------------|------------|--------------|-----------|------------|--------|--------|---|------|---|
| Query Match | 1.3%       | Score        | 63.6      | DB 1       | Length | 1830   |   |      |   |
| Best Local  | Similarity | 65.5%        | Pred. No. | 1.4e-08    |        |        |   |      |   |
| Matches     | 93         | Conservative | 0         | Mismatches | 49     | Indels | 0 | Gaps | 0 |

RESULT 11  
US-08-858-311-2  
; Sequence 2, Application US/08858311-  
; Patent No. 5876939

```

      FILING DATE: 14-JUN-1994
      ATTORNEY/AGENT INFORMATION:
      NAME: Campbell, Cathryn
      REGISTRATION NUMBER: 31,815
      REFERENCE/DOCKET NUMBER: P-1J 1389
      TELECOMMUNICATION INFORMATION:
      TELEPHONE: (619) 535-9001
      TELEFAX: (619) 535-8949
      INFORMATION FOR SEQ ID NO: 2:
      SEQUENCE CHARACTERISTICS:
      LENGTH: 1830 base pairs
      TYPE: nucleic acid
      STRANDEDNESS: double
      TOPOLOGY: linear
      MOLECULE TYPE: cDNA
      US-08-858-311-2

Query Match      1.3%; Score 63.6; DB 2; Length 1830;
Best Local Similarity 65.5%; Pred. No. 1.4e-08;
Matches 93; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

Db 4229 cattatgttaaacagcgtttgcgaagggagcagcctctgaaagcggcgtctgaaag 4298
      |||||
Db 363 CATTTATGTGAAGGCTGTATTCCCGAGGAGCACAGCTGTGATGTAGTAATTCACAA 422
      |||||
Qy 4299 gggcgcatacatcattgcttgcatacgggcagagctagaagagatccaccatgaagaagc 4358
      |||||
Db 423 AGGTGATGCGCGTCCTACGCTGATCAATGAGATTAGTGTAGAGAGAGCACCACCATTAAGCAAGC 482
      |||||
Qy 4359 tgttgccatccttaacggaca 4380
      |||||
Db 483 TGTGGAACACTGAGAAATAC 504
      |||||

RESULT 12
US-09-151-611-2
Sequence 2, Application US/09151611
Patent No. 5958731
GENERAL INFORMATION:
APPLICANT: Yu, Henry
APPLICANT: Au-Young, Janice
APPLICANT: Paterson, Chandra
TITLE OF INVENTION: CELL JUNCTION PDZ PROTEIN
FILE REFERENCE: PR-0559 US
CURRENT APPLICATION NUMBER: US/09/151,611
CURRENT FILING DATE: 1998-09-11
NUMBER OF SEQ ID NOS: 3
SOFTWARE: PERL Program
SEQ ID NO 2
LENGTH: 1396
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: 1974337
US-09-151-611-2

Query Match      1.1%; Score 56; DB 2; Length 1396;
Best Local Similarity 57.4%; Pred. No. 2.2e-06;
Matches 101; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

Db 4230 agacttaccattatgtttaaacagcgtttgcgaagggagcagcctctgaagcggagc 4289
      |||||
Db 749 aaattcccccattatatactctgcataatctcctggaagggtggttgaaagacaggaag 808
      |||||
Qy 4290 tctgaagaaggcgatcagatcatgtctgtcaatggtggcagaagcttagaagaagtcacca 4349
      |||||
Db 809 cctcaaaaggagaaacacgcgtgctatcgtgaacggaagtgtgtgtggaagaggaacacca 868
      |||||
Qy 4350 tgaagaagctgttcacatccttaaacgagacaagaagcagctgtcaccttgaatgttc 4405
      |||||
Db 869 tgaagaagctgttgaaactacacacagcgtctcaaaagacagcgtcaagcgtgtgtgc 924
      |||||

```



## RESULT 13

US-09-276-531-77  
; Sequence 77, Application US/09276531  
; Patent No. 6183968

## GENERAL INFORMATION:

APPLICANT: Bandman, Olga  
APPLICANT: Lal, Preeti  
APPLICANT: Hillman, Jennifer L.  
APPLICANT: Yue, Henry  
APPLICANT: Reddy, Roopa  
APPLICANT: Guegler, Karl J.  
APPLICANT: Baughn, Mariah R.  
TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF GENES ENCODING  
TITLE OF INVENTION: RECEPTORS AND PROTEINS ASSOCIATED WITH CELL PROLIFERATION  
NUMBER OF SEQUENCES: 134  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
STREET: 3174 PORTER DRIVE  
CITY: PALO ALTO  
STATE: CALIFORNIA  
COUNTRY: USA  
ZIP: 94304

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/276,531  
FILING DATE: Herewith

## CLASSIFICATION:

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/079,677  
FILING DATE: March 27, 1998

## CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Lynn E. Murty, Ph.D.

REGISTRATION NUMBER: 42,918

REFERENCE/DOCKET NUMBER: PA-0008 US

TELECOMMUNICATION INFORMATION:

TELEPHONE: (650) 855-0555

TELEFAX: (650) 845-4166

INFORMATION FOR SEQ ID NO: 77:

SEQUENCE CHARACTERISTICS:

LENGTH: 1042 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

IMMEDIATE SOURCE:

LIBRARY: BLADYUT04

CLONE: 1555118

US-09-276-531-77

Query Match

Best Local Similarity 59.2%; Score 52.8; DB 4; Length 1042;

Matches 90; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

Db 16 AGACTGGGATTTTATCAATCCATCTTCATGAGGCGCTCTTTTAAAGATGCTCG 75

Qy 4230 agactaccattatgttaaacagtgcttgaaggagagcagcctctgaaggagagc 4289

Db 4290 tctgaagaaggcgatcagatcatctgttcaatggcgagagctagaaggagtcacca 4349

Qy 76 TCTGCAATGAATGACACGCTGATTCAGATTATGGGAATCTCTTTGGGAAGTCCAA 135

Qy 4350 tgaagaagctgtgtccatctccttaaacgagaca 4381

Db 136 CCACGAAGCTATGGAAACCTTGGGCGGTCAA 167

## RESULT 14

US-08-580-680-2  
; Sequence 2, Application US/08580680  
; Patent No. 5766866

## GENERAL INFORMATION:

APPLICANT: Center, David M.  
APPLICANT: Cruikshank, William W.  
APPLICANT: Kornfeld, Hardy  
TITLE OF INVENTION: LYMPHOCYTE CHEMOTACTANT FACTOR AND  
TITLE OF INVENTION: USES THEREOF  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Scully, Scott, Murphy & Presser  
STREET: 400 Garden City Plaza  
CITY: Garden City  
STATE: New York  
COUNTRY: USA  
ZIP: 11530-0299

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/580,680  
FILING DATE: 29-DEC-1995

## CLASSIFICATION:

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/480,156

FILING DATE: 07-JUN-1995

ATTORNEY/AGENT INFORMATION:

NAME: Frank S. Digioi

REGISTRATION NUMBER: 31,346

REFERENCE/DOCKET NUMBER: 9874ZY

TELECOMMUNICATION INFORMATION:

TELEPHONE: (516) 742-4343

TELEFAX: (516) 742-4366

TELEX: 230 901 SANS UR

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 2150 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

US-08-580-680-2

Query Match

Best Local Similarity 61.1%; Score 49.4; DB 1; Length 2150;

Matches 80; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

Db 592 ATTACGTTACAGAGTGTTCATTAATGGGCTGCTCCAGGAAGGACTATTCGAAG 651

Qy 4240 attatgttaaacagtgcttgaaggagagcagcctctgaaggagagcgtctgaagg 4299

Db 4300 ggcgcatcagatcgtcttcaatggcgagagctagaaggagtcaccaatgaaga 4359

Qy 652 GGCATGAATGATTTTCATCAACGCGCAAGTCTCAAGGGGACGACCATGATGCC 711

Db 4360 gtgcgcatcct 4370

Qy 712 TTGGCCATCCT 722

Db 712 TTGGCCATCCT 722

Query Match

Best Local Similarity 61.1%; Score 49.4; DB 1; Length 2150;

Matches 80; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

Db 592 ATTACGTTACAGAGTGTTCATTAATGGGCTGCTCCAGGAAGGACTATTCGAAG 651

Qy 4240 attatgttaaacagtgcttgaaggagagcagcctctgaaggagagcgtctgaagg 4299

Db 4300 ggcgcatcagatcgtcttcaatggcgagagctagaaggagtcaccaatgaaga 4359

Qy 652 GGCATGAATGATTTTCATCAACGCGCAAGTCTCAAGGGGACGACCATGATGCC 711

Db 4360 gtgcgcatcct 4370

Qy 712 TTGGCCATCCT 722

Db 712 TTGGCCATCCT 722

TITLE OF INVENTION: LYMPHOCYTE CHEMOATTRACTANT FACTOR AND  
 TITLE OF INVENTION: USES THEREOF  
 NUMBER OF SEQUENCES: 2  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Scully, Scott, Murphy & Presser  
 STREET: 400 Garden City Plaza  
 CITY: Garden City  
 STATE: New York  
 COUNTRY: USA  
 ZIP: 11530-0299  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patentin Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/480,156A  
 FILING DATE: 07-JUN-1995  
 CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Frank S. DIGLIO  
 REGISTRATION NUMBER: 31,346  
 REFERENCE/DOCKET NUMBER: 98742Y  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (516) 742-4343  
 TELEFAX: (516) 742-4366  
 TELE: 230 901 SANS UR  
 INFORMATION FOR SEQ ID NO: 2:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 2150 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: DNA (genomic)  
 US-08-480-156A-2

Query Match 1.0%; Score 49.4; DB 1; Length 2150;  
 Best Local Similarity 61.1%; Pred. No. 0.00026;  
 Matches 80; Conservative 0; Mismatches 51; Indels 0; Gaps 0;  
 QY 4240 attatgttaaacagctgttgcacaaggagcagcctctgaagaagcagctctgaaaagg 4239  
 DB 592 ATTACGGTTTCACAGAGTGTTCCTCAAGTGGCTGCCAGGAAGGACTATTTCAGAG 651  
 QY 4300 ggcagtcagatcattgctcaatgagcagagcttaagaaggagtcacccatgaagaagct 4359  
 DB 652 GGCATATGAGCTTCTTCATCAACGCAAGTCTCTCAAGGGGACCAACGACCATGATGCC 711  
 QY 4360 gtgcacatcct 4370  
 DB 712 TTGGCCATCTT 722

Search completed: July 12, 2001, 21:46:52  
 Job time: 10205 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 12, 2001, 17:37:37 ; Search time 3448.6 Seconds  
(without alignments)  
13376.421 Million cell updates/sec

Title: US-09-502-698-3

Perfect score: 4880  
Sequence: 1 cccggccccggcgacagtg.....tctctcccccccccaact 4880

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 10228115 seqs, 4726426750 residues

Total number of hits satisfying chosen parameters: 20456230

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

EST: \*  
1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_est3:\*  
4: gb\_est4:\*  
5: gb\_est5:\*  
6: gb\_est6:\*  
7: gb\_est7:\*  
8: gb\_est8:\*  
9: gb\_est9:\*  
10: gb\_est10:\*  
11: gb\_est11:\*  
12: gb\_est12:\*  
13: gb\_est13:\*  
14: gb\_est14:\*  
15: gb\_est15:\*  
16: gb\_est16:\*  
17: gb\_est17:\*  
18: gb\_est18:\*  
19: gb\_est19:\*  
20: gb\_est20:\*  
21: gb\_est21:\*  
22: gb\_est22:\*  
23: gb\_est23:\*  
24: gb\_est24:\*  
25: gb\_est25:\*  
26: gb\_est26:\*  
27: gb\_est27:\*  
28: gb\_est28:\*  
29: gb\_est29:\*  
30: gb\_est30:\*  
31: gb\_est31:\*  
32: gb\_est32:\*  
33: gb\_est33:\*  
34: gb\_est34:\*  
35: gb\_est35:\*  
36: gb\_est36:\*  
37: gb\_est37:\*  
38: gb\_est38:\*  
39: gb\_est39:\*  
40: gb\_est40:\*  
41: gb\_est41:\*  
42: gb\_est42:\*  
43: gb\_est43:\*

44: em\_esthum10:\*  
45: em\_esthum11:\*  
46: em\_esthum12:\*  
47: em\_esthum13:\*  
48: em\_esthum14:\*  
49: em\_esthum15:\*  
50: em\_esthum16:\*  
51: em\_esthum17:\*  
52: em\_esthum18:\*  
53: em\_esthum19:\*  
54: em\_esthum20:\*  
55: em\_esthum21:\*  
56: em\_esthum22:\*  
57: em\_esthum23:\*  
58: em\_esthum24:\*  
59: em\_esthum25:\*  
60: em\_esthum26:\*  
61: em\_esthum27:\*  
62: em\_esthum28:\*  
63: em\_esthum29:\*  
64: em\_esthum30:\*  
65: em\_esthum31:\*  
66: em\_esthum32:\*  
67: em\_esthum33:\*  
68: em\_esthum34:\*  
69: em\_esthum35:\*  
70: em\_esthum36:\*  
71: em\_esthum37:\*  
72: em\_esthum38:\*  
73: em\_esthum39:\*  
74: em\_esthum40:\*  
75: em\_esthum41:\*  
76: em\_esthum42:\*  
77: em\_esthum43:\*  
78: em\_esthum44:\*  
79: em\_esthum45:\*  
80: em\_esthum46:\*  
81: em\_esthum47:\*  
82: em\_esthum48:\*  
83: em\_esthum49:\*  
84: em\_esthum50:\*  
85: em\_esthum51:\*  
86: em\_esthum52:\*  
87: em\_esthum53:\*  
88: em\_esthum54:\*  
89: em\_esthum55:\*  
90: em\_esthum56:\*  
91: em\_esthum57:\*  
92: em\_esthum58:\*  
93: em\_esthum59:\*  
94: em\_esthum60:\*  
95: em\_esthum61:\*  
96: em\_esthum62:\*  
97: em\_esthum63:\*  
98: em\_esthum64:\*  
99: em\_esthum65:\*  
100: em\_esthum66:\*  
101: em\_esthum67:\*  
102: em\_esthum68:\*  
103: em\_esthum69:\*  
104: em\_esthum70:\*  
105: em\_esthum71:\*  
106: em\_esthum72:\*  
107: em\_esthum73:\*  
108: em\_esthum74:\*  
109: em\_esthum75:\*  
110: em\_esthum76:\*  
111: em\_esthum77:\*  
112: em\_esthum78:\*  
113: em\_esthum79:\*  
114: em\_esthum80:\*  
115: em\_esthum81:\*  
116: em\_esthum82:\*

117: gb\_est48:\*  
 118: gb\_est49:\*  
 119: gb\_est50:\*  
 120: gb\_est51:\*  
 121: gb\_est52:\*  
 122: gb\_est53:\*  
 123: gb\_est54:\*  
 124: gb\_est55:\*  
 125: gb\_est56:\*  
 126: gb\_est57:\*  
 127: gb\_est58:\*  
 128: gb\_est59:\*  
 129: gb\_est60:\*  
 130: gb\_est61:\*  
 131: gb\_est62:\*  
 132: gb\_est63:\*  
 133: gb\_est64:\*  
 134: gb\_est65:\*  
 135: gb\_est66:\*  
 136: gb\_est67:\*  
 137: gb\_est68:\*  
 138: gb\_est69:\*  
 139: gb\_est70:\*  
 140: gb\_est71:\*  
 141: gb\_est72:\*  
 142: gb\_est73:\*  
 143: gb\_est74:\*  
 144: gb\_est75:\*  
 145: gb\_est76:\*  
 146: gb\_est77:\*  
 147: gb\_est78:\*  
 148: gb\_est79:\*  
 149: gb\_est80:\*  
 150: gb\_est81:\*  
 151: gb\_est82:\*  
 152: gb\_est83:\*  
 153: gb\_est84:\*  
 154: gb\_est85:\*  
 155: gb\_est86:\*  
 156: gb\_est87:\*  
 157: gb\_est88:\*  
 158: gb\_est89:\*  
 159: gb\_est90:\*  
 160: gb\_est91:\*  
 161: gb\_est92:\*  
 162: gb\_est93:\*  
 163: gb\_est94:\*  
 164: gb\_est95:\*  
 165: gb\_est96:\*  
 166: gb\_est97:\*  
 167: gb\_est98:\*  
 168: gb\_est99:\*  
 169: gb\_est100:\*  
 170: gb\_est101:\*  
 171: gb\_est102:\*  
 172: gb\_est103:\*  
 173: gb\_est104:\*  
 174: gb\_est105:\*  
 175: gb\_est106:\*  
 176: gb\_est107:\*  
 177: gb\_est108:\*  
 178: gb\_est109:\*  
 179: gb\_est110:\*  
 180: gb\_est111:\*  
 181: gb\_est112:\*  
 182: gb\_est113:\*  
 183: gb\_est114:\*  
 184: gb\_est115:\*  
 185: gb\_est116:\*  
 186: gb\_est117:\*  
 187: gb\_est118:\*  
 188: gb\_est119:\*  
 189: gb\_est120:\*

190: gb\_est121:\*  
 191: gb\_est122:\*  
 192: gb\_est123:\*  
 193: gb\_est124:\*  
 194: gb\_est125:\*  
 195: gb\_est126:\*  
 196: gb\_est127:\*  
 197: gb\_est128:\*  
 198: gb\_est129:\*  
 199: gb\_est130:\*  
 200: gb\_est131:\*  
 201: gb\_est132:\*  
 202: gb\_est133:\*  
 203: gb\_est134:\*  
 204: gb\_est135:\*  
 205: gb\_est136:\*  
 206: gb\_est137:\*  
 207: gb\_est138:\*  
 208: gb\_est139:\*  
 209: gb\_est140:\*  
 210: gb\_est141:\*  
 211: gb\_est142:\*  
 212: gb\_est143:\*  
 213: gb\_est144:\*  
 214: gb\_est145:\*  
 215: gb\_est146:\*  
 216: gb\_est147:\*  
 217: gb\_est148:\*  
 218: gb\_est149:\*  
 219: gb\_est150:\*  
 220: gb\_est151:\*  
 221: gb\_est152:\*  
 222: gb\_est153:\*  
 223: gb\_est154:\*  
 224: gb\_est155:\*  
 225: gb\_est156:\*  
 226: gb\_est157:\*  
 227: gb\_est158:\*  
 228: gb\_est159:\*  
 229: gb\_est160:\*  
 230: gb\_est161:\*  
 231: gb\_est162:\*  
 232: gb\_est163:\*  
 233: gb\_est164:\*  
 234: gb\_est165:\*  
 235: gb\_est166:\*  
 236: gb\_est167:\*  
 237: gb\_est168:\*  
 238: gb\_est169:\*  
 239: gb\_est170:\*  
 240: gb\_est171:\*  
 241: gb\_est172:\*  
 242: gb\_est173:\*  
 243: gb\_est174:\*  
 244: gb\_est175:\*  
 245: gb\_est176:\*  
 246: gb\_est177:\*  
 247: gb\_est178:\*  
 248: gb\_est179:\*  
 249: gb\_est180:\*  
 250: gb\_est181:\*  
 251: gb\_est182:\*  
 252: gb\_est183:\*  
 253: gb\_est184:\*  
 254: gb\_est185:\*  
 255: gb\_est186:\*  
 256: gb\_est187:\*  
 257: gb\_est188:\*  
 258: gb\_est189:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## JOURNAL REFERENCE

[illegible]

```

Rot = 37.5. Second strand cDNA was prepared with the primer adapted
of sequence [5',
GAGAGGAGCATTCCTCAGTTAATTAAATTAATCCCCCCCCCCC 3']. cDNA was cleaved
with XhoI and SstI. Cloning sites, 5' end: XhoI; 3' end: SstI.
Host: 501R.

FEATURES
        source
            location/Qualifiers
                1..2756
                    /organism="Mus musculus"
                    /strain="C57BL/6J"
                    /db_xref="taxon:10090"
                    /db_xref="MGI:1343489"
                    /db_xref="MGI:1912867"
                    /clone="2610029H06"
                    /clone_1lb="Riken full-length enriched mouse cDNA library"
                    /dev_stage="10 days embryo"
                    <1..1042
                    /note="putative"

```



```

Qy 4691 gtacatgggcaagaattatcatgctaagctggttagtataaagaataattctaa 4750
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1878 gtcacttgacacaa-aattatattatgttcattctttagctgcaaaacaaatgatac 1936
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 4751 agctaaccaagaagaatggtcagtaagttagatgaaatgaaataataaataaag 4810
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1937 agttagcgaattaggttaagatgaaatttacaataaagaacagattttcaggttta 1996
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 4811 aagaataatc 4820
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1997 taggacttct 2006
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 2
AV691714 632 bp mRNA EST 25-SEP-2000
LOCUS AV691714 GKC Homo sapiens cDNA clone GKCM01 5', mRNA sequence.
DEFINITION AV691714
ACCESSION AV691714
VERSION AV691714.1 GI:10293577
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 632)
AUTHORS Wu,T., Qian,B., Huang,Q., Kang,B., Gao,X., Xu,Z., Xiao,H.,
XU,X., LI,N., Peng,Y., Liu,F., Qu,J., Song,H., Cheng,Z., Zeng,L.,
Xu,S., Gu,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M., Lu,G., Yang
Y., Gu,Y., Chen,Z. and Han,Z.
Homo sapiens cDNA GK-clones
Unpublished (2000)
JOURNAL Contact: Zeguang Han
COMMENT Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919(ex.45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.

FEATURES
source
1..632
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="GKCM01"
/clone_1lb="GKC"
/tissue_type="hepatocellular carcinoma"
/dev_stage="Adult"
/lab_host="SOLR"
/note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2:
XhoI"

BASE COUNT 191 a 118 c 179 g 139 t 5 others
ORIGIN

Query Match 12.0%; Score 586.8; DB 32; Length 632;
Best Local Similarity 97.4%; Pred. No. 1,le-148;
Matches 605; Conservative 0; Mismatches 12; Indels 4; Gaps 1;

Qy 3171 tgaagttatgaagaagcagcatgtaagaagccttggtcgagatcagat 3330
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 tgaagttatgaagaagcagcatgtaagaagccttggtcgagatcagat 60
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 3231 cttaagagtgaaatgacttgacttggaagcagcacatgatgaatcaatctct 3290
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 cttaagagtgaaatgacttgacttggaagcagcacatgatgaatcaatctct 120
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 3291 ggaacagagccacagaagtgctgacacactacagagatgaagccacatacaaga 3350
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 121 ggaacagagccacagaagtgctgacacactacagagatgaagccacatacaaga 180
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 3351 ggaagagtgctgacacactacatctgagctgcagaagaagccggaaagcctagg 3410
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 181 ggaggaagtgctgacacactacatctgagctgcagaagaagccggaaagcctagg 240
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

Qy 3411 attaatattgtgtgtaaaagaacgatctagatattgtgtcagacattgtcaagg 3470
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 241 attaatattgtgtgtgtaaaagaacgatctagatattgtgtcagacattgtcaagg 296
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 3471 aggaattcagatcccgatggaagactgacccggagagacacatacttggtaagtg 3530
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 297 aggaattcagatcccgatggaagactgacccggagagacacatacttggtaagtg 356
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 3531 ggaagaagctgtgtaatgctccccaagaagcgtgtgcgccttgctaaagtgtccctagg 3590
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 357 ggaagaagctgtgtaatgctccccaagaagcgtgtgcgccttgctaaagtgtccctagg 416
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 3591 caagaataacctggaagtggaagaatcaagctgtgtcattcattcagaagagggcc 3650
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 417 caagaataacctggaagtggaagaatcaagctgtgtcattcattcagaagagggcc 476
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 3651 atctcaacagccaggtgagtgaaagcagccgtctcttctcacttctcactcctg 3710
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 477 atctcaacagccaggtgagtgaaagcagccgtctcttctcacttctcactcctg 536
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 3711 atccagtaacatctgagctcagtgaaagtgctcaaaagaatgcatgcatctgaaat 3770
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 537 atccagtaacatctgagctcagtgaaagtgctcaaaagaatgcatgcatctgaaat 596
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 3771 acagagattaagaacagctga 3791
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 597 acagagattaagaacagctga 617
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 3
AV695000 634 bp mRNA EST 25-SEP-2000
LOCUS AV695000 GKC Homo sapiens cDNA clone GKCM01 5', mRNA sequence.
DEFINITION AV695000
ACCESSION AV695000
VERSION AV695000.1 GI:10296863
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 634)
AUTHORS Wu,T., Qian,B., Huang,Q., Kang,B., Gao,X., Xu,Z., Xiao,H.,
XU,X., LI,N., Peng,Y., Liu,F., Qu,J., Song,H., Cheng,Z., Zeng,L.,
Xu,S., Gu,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M., Lu,G., Yang
Y., Gu,Y., Chen,Z. and Han,Z.
Homo sapiens cDNA GK-clones
Unpublished (2000)
JOURNAL Contact: Zeguang Han
COMMENT Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919(ex.45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.

FEATURES
source
1..634
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="GKCM01"
/clone_1lb="GKC"
/tissue_type="hepatocellular carcinoma"
/dev_stage="Adult"
/lab_host="SOLR"
/note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2:
XhoI"

BASE COUNT 193 a 120 c 179 g 140 t 2 others
ORIGIN

Query Match 11.9%; Score 581.2; DB 32; Length 634;
Best Local Similarity 97.7%; Pred. No. 3,8e-147;

```





QY 3411 attaaagtattgtttggtgtaaaagaagaacactagctattgtttcagacattgcaaaagg 3470  
Db 241 ATTTAAGTATTGTTGGT----GAACAGATACCTGGAGTATTGTGTCAGACATTGCCAAAG 296

```

Oy 3471 agaatcgcagatcccgatggaagactgatacaggagaccagatatattggtgaatg 3530
|||||
Db 297 AGGAATTGCATGCCGATGTAAGATGATGCAGGAGACAGATATTAAATGTAATGG 356
|||||
Oy 3531 ggaagaagtcgttaataatgctccccaagaagcggttcgcgcttgcctaaagtctccctagg 3590
|||||
Db 357 GGAAGAGCTTGTGTAATGTCACACCAAGAACGCGTTGCCGCTTGTAAAGTGTCCCTAAG 416
|||||
Oy 3591 cacagtaacctgtgaagttggaagaatcaaaagctgctccatctcatctcaagagaagc 3650
|||||
Db 417 CACAGTAACCTTGGAAGTGAAGATCAAAAGCTGTCATCTTCATTCACAGAGAGAGGCC 476
|||||
Oy 3651 atctcaaacaccagcaggtagtgaagcagcctgtctctcttcaacttccacactctg 3710
|||||
Db 477 ATCTCAAGCAGCCAGGAGTGAAGGAGCGCTGTCACTTTTCACCTTCCACTCTGG 536
|||||
Oy 3711 atccagtaacatctgtagtcaactctggaagtagtcaagaagaatgcatctgcatctgaat 3770
|||||
Db 537 ATCCAGTACATCTGATGCTACCTGGAAGAGCTCAAGAAAGATGCA-TGGCATCTGAAT 595
|||||
Oy 3771 acaggagtaagaa 3784
|||||
Db 596 ACAGGATTAAGCA 609
|||||

RESULT 7
AV696967 615 bp mRNA EST 25-SEP-2000
LOCUS AV696967 GKC Homo sapiens cDNA clone GKCNF08 5', mRNA sequence.
DEFINITION AV696967.1 GI:10298830
ACCESSION AV696967.1
VERSION AV696967.1
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 615)
AUTHORS Wu, T., Qian, B., Huang, Q., Kang, B., Gao, X., Xu, Z., Xiao, H.,
Xu, X., Li, N., Peng, Y., Liu, F., Qu, J., Song, H., Cheng, Z., Zeng, L.,
Xu, S., Gu, W., Tu, Y., Jia, J., Fu, G., Ren, S., Zhong, M., Lu, G., Yang
Y., Gu, Y., Chen, Z., and Han, Z.
Homo sapiens cDNA GK- clones
Unpublished (2000)
JOURNAL Contact: Zeguang Han
COMMENT Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919(ex.45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.

FEATURES
source
1..615
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="GKCNF08"
/clone_id="GKC"
/tissue_type="hepatocellular carcinoma"
/dev_stage="Adult"
/lab_host="SOLR"
/note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2:
XhoI"

BASE COUNT 191 a 115 c 174 g 134 t 1 others
ORIGIN
Query Match 11.7%; Score 570.2; DB 32; Length 615;
Best Local Similarity 97.7%; Pired. NO. 3.7e-14;
Matches 600; Conservative 0; Mismatches 9; Indels 5; Caps 2;

Oy 3711 tgaagttatgaagaagagagcatgtlaaagatggaagactcgggctgagatcagat 3730
|||||
Db 1 TGAAGTTATGAAGAAGAGAGCATGTAAAGATGGAAGACTGTGGCTGGAGATCAGAT 60
|||||

```

```

Oy 3331 cttagagtgatgaatctgacttgagaagccacacatgatgaacatcaatgctc 3290
|||||
Db 61 CTTAGAGTGAATGATGATGACTTGTAGAAAGCCACACATCATATACCATATGTCCT 120
|||||
Oy 3391 gagaacagcgcacagagagtgccctgcaactctacagagatgagcccatcaaga 3350
|||||
Db 121 GAGACACAGCCACAGAGAGTGCCTGACACTCTACAGATGAGGCCCATCAAGA 180
|||||
Oy 3351 ggaagaagtggtgacacccctcactatgtgctgcagaagaagccggaaagcctagg 3410
|||||
Db 181 GGAGAAAGTGTGTACACCCCTCACTATTGAGCTCACAAGAGCCGGAAAGGCTTAGG 240
|||||
Oy 3411 attagatctgtgtlaaagaacagatcactgagatttctgcagacatgtcagaag 3470
|||||
Db 241 ATTAAGTATTGTTGCT----GAAAGTACTGAGATTTGTCACACATGTCAAGG 296
|||||
Oy 3471 agaatcgcagatcccgatggaagactgatacaggagaccagatatattggtgaatg 3530
|||||
Db 297 AGGAATTGCATGCCGATGTAAGATGATGCAGGAGACAGATATTAAATGTAATGG 356
|||||
Oy 3531 ggaagaagtcgttaataatgctccccaagaagcggttcgcgcttgcctaaagtctccctagg 3590
|||||
Db 357 GGAAGAGCTTGTGTAATGTCACACCAAGAACGCGTTGCCGCTTGTAAAGTGTCCCTAAG 416
|||||
Oy 3591 cacagtaacctgtgaagttggaagaatcaaaagctgctccatctcatctcaagagaagc 3650
|||||
Db 417 CACAGTAA-CTTGAAGTGAAGATCAAAAGCTGTCATCTTCATTCACAGAGAGAGGCC 475
|||||
Oy 3651 atctcaaacaccagcaggtagtgaagcagcctgtctctcttcaacttccacactctg 3710
|||||
Db 476 ATCTCAAGCAGCCAGGAGTGAAGGAGCGCTGTCACTTTTCACCTTCCACTCTGG 535
|||||
Oy 3711 atccagtaacatctgtagtcaactctggaagtagtcaagaagaatgcatctgcatctgaat 3770
|||||
Db 536 ATCCAGTACATCTGATGCTACCTGGAAGAGCTCAAGAAAGATGCAATGGCATCTGAAT 595
|||||
Oy 3771 acaaggatlaagaa 3784
|||||
Db 596 ACAGGATTAAGCA 609
|||||

RESULT 8
AV690061 622 bp mRNA EST 25-SEP-2000
LOCUS AV690061 GKC Homo sapiens cDNA clone GKCAN09 5', mRNA sequence.
DEFINITION AV690061
ACCESSION AV690061
VERSION AV690061.1 GI:10291924
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 622)
AUTHORS Wu, T., Qian, B., Huang, Q., Kang, B., Gao, X., Xu, Z., Xiao, H.,
Xu, X., Li, N., Peng, Y., Liu, F., Qu, J., Song, H., Cheng, Z., Zeng, L.,
Xu, S., Gu, W., Tu, Y., Jia, J., Fu, G., Ren, S., Zhong, M., Lu, G., Yang
Y., Gu, Y., Chen, Z., and Han, Z.
Homo sapiens cDNA GK- clones
Unpublished (2000)
JOURNAL Contact: Zeguang Han
COMMENT Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919(ex.45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.

FEATURES
source
1..622
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="GKCAN09"

```

|                           |   |
|---------------------------|---|
| AUTHORS                   | HILLIER, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S., Krizman, D., Kucada, T., Lacey, M., Le, N., Lennon, G., Maiza, M., Martin, J., Moore, B., Schellenberg, K., Stepien, M., Tan, F., Theising, B., White, Y., Wylie, T., Waterston, R. and Wilson, R.   |
| TITLE                     | WashU-NCI human EST Project   |
| JOURNAL                   | Unpublished (1997)  |
| COMMENT                   | Contact: Wilson RK<br>Washington University School of Medicine<br>444 Forest Park Parkway, Box 8501, St. Louis, MO 63108<br>Tel: 314 286 1800<br>Fax: 314 286 1810<br>Email: est@wustl.wustl.edu<br>This clone is available royalty-free through LNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.<br>Seq primer: -40RP from Gldco<br>High quality sequence stop: 424.  |
| FEATURES                  | Location/Qualifiers<br>1. 582<br>/organism="Homo sapiens"<br>/db_xref="taxon:9606"<br>/clone="IMAGE:2518539"<br>/clone_lib="Schneider fetal brain 00004"<br>/sex="male"<br>/tissue_type="frontal lobe"<br>/dev_stage="5 months post-conception"<br>/lab_host="DH10B"<br>/note="Organ: brain; Vector: pBluescript SK (Stratagene); Site_1: SstI; Site_2: XhoI; Double-stranded cDNA was prepared from human fetal brain tissue. 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence:<br>5'-GAGAGAGAGAAGACGCTCAGAGATCCTTAATTAATTAATCCCCCCCCC-3' and 3' adaptor sequence:<br>5'-GAGAGAGAGACTCGAGATTTTCTTTTCTTTT-3'. The library was size-selected for >0.5 kb inserts and has an average insert size estimated at 1.2 kb. This library was constructed using the CAP-trapper method for full-length enrichment and has not undergone amplification. Library was constructed by Dr. Claudio Schneider (LNCIB-Area Science Park, Trieste, Italy)." |
| BASE COUNT                | 161 a 115 c 129 g 176 t 1 others  |
| ORIGIN                    |   |
| Query Match               | 11.7%; Score 569; DB 103; Length 582;   |
| Best Local Similarity     | 99.7%; Pred. NO. 7.8e-144;  |
| Matches 580; Conservative | 0; Mismatches 1; Indels 1; Gaps 1;  |
| OY 4123                   | atattccagatgatttgagaccctcccaatgtaagctatatacactagagcgagacca 4182   |
| Db 1                      | ATATTTCAGAGATGATTAGGACCTCCTCAATGTAGTATACACTGAGCGAGACCA 60   |
| OY 4183                   | gatgcttagagcttcagatagtttggaagatattgcaagccctcagagactaccatt 4242  |
| Db 61                     | GATGGCTTAGGCTTGAGTATAGTTGAGAGATATGGACCCCTCAATGAGACTTACCATT 120  |
| OY 4243                   | tatgttaaaacagtgcttgcgaaggagcagcctctgaaagcagcgcctgaaaggggc 4302  |
| Db 121                    | TATGTAAACAGCTTTTGCAAAAGGAGACAGCCTCTGAAAGAGGAGCGCTGAAAAAGGGC 180   |
| OY 4303                   | gatcagatcatgtctcaatgtygcagagctctagaaggagtcaccactgaaagacgtgt 4362  |
| Db 181                    | GATCAGATCATGTGCTCAATGGCGAGAGTGTAGAAAGAGTCAACCATCAATCAACACTGTT 240   |
| OY 4363                   | ggcaatccctaaacgaaaggacgtctcaacttgatgagtgctctcttgaaattggcgc 4422   |
| Db 241                    | GCCATCCCTTAAACGGAACAAAGGCATCTGCATTGTGAATGGTCTCTTGAATTGGCTGC 300   |
| OY 4423                   | cagaattgaacacccaaacccctagaccacctcactctgtaaaagagaatgcaatgctgc 4482   |
| Db 301                    | CAGAAATTGACCAACCCCAACCCCTTAGCTCACCCTCTTGTGAAGAGAAATGCACTGCTCC 360   |
| OY 4483                   | tgacaatttatatgctgtgttcagccgggctctcaaaactgtaagggggaaataaactt 4542  |

| DB                        | 361   | TCACATATTTTATGCGTGTTCAGCCGGGTCTTCAAAACTGTAGGGGGGAAATTAACACTT  | 420                 |
|---------------------------|---|---|---------------------|
| OY                        | 4543  | aagttcttttctcaatcagaatagcttctcttacttgcaacccaatcattttctt       | 4602                |
| Db                        | 421   | AAGTTCTTTTCTCATCTAGAAAGCTTCTCTTACTGTCAACCTAACTATTTTCTT        | 480                 |
| OY                        | 4603  | tctctctgcatctgtgacttaagaagaagaatattgtgtggtg-aaatcggtt         | 4661                |
| Db                        | 481   | TTCCTCTTGCAATNTGTGACTTAAGAGAACAGATATTGTGTAGGTAAATTCGT         | 540                 |
| OY                        | 4662  | ttattctgaggaataltcctaagttcttgatgtacatgggcaa                   | 4703                |
| Db                        | 541   | TTATTTGTGGAGATATCTAATGTTTGTAGTCACATGGCAA                      | 582                 |
| RESULT                    | 10  |   |                     |
| LOCUS                     | Bf675982  | 855 bp  | mRNA                |
| DEFINITION                | 6020839339p1 NIH_MGC_83 Homo sapiens CDNA clone IMAGE:4248369     | 5'  | EST 21-DEC-2000     |
| ACCESSION                 | Bf675982  |   |                     |
| VERSION                   | Bf675982.1  | GI:11949877   |                     |
| KEYWORDS                  | EST.  |   |                     |
| SOURCE                    | human.  |   |                     |
| ORGANISM                  | Homo sapiens  |   |                     |
| REFERENCE                 | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; |   |                     |
| AUTHORS                   | Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.         |   |                     |
| TITLE                     | NIH-MGC http://mgc.ncl.nih.gov/.                                  |   |                     |
| JOURNAL                   | National Institutes of Health, Mammalian Gene Collection (MGC)    |   |                     |
| COMMENT                   | Unpublished (1999)  |   |                     |
|                           | Contact: Robert Strausberg, Ph.D.                                 |   |                     |
|                           | Email: cga@b-remail.nih.gov                                       |   |                     |
|                           | Tissue Procurement: CLONETECH Laboratories, Inc.                  |   |                     |
|                           | CDNA Library Preparation: CLONETECH Laboratories, Inc.            |   |                     |
|                           | CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)         |   |                     |
|                           | DNA Sequencing by: Incyte Genomics, Inc.                          |   |                     |
|                           | Clone distribution: MGC clone distribution information can be     |   |                     |
|                           | found through the I.M.A.G.E. Consortium/LLNL at:                  |   |                     |
|                           | http://image.llnl.gov   |   |                     |
|                           | Plate: L1CML066 row: m column: 10                                 |   |                     |
|                           | High quality sequence stop: 614.                                  |   |                     |
| FEATURES                  |   |   |                     |
| SOURCE                    | 1..855  |   |                     |
|                           | /organism="Homo sapiens"  |   |                     |
|                           | /db_xref="taxon:9606"   |   |                     |
|                           | /clone="IMAGE:4248369"  |   |                     |
|                           | /clone_11b="NIH_MGC_83"   |   |                     |
|                           | /lab_host="DH10B (TI phage-resistant)"                            |   |                     |
|                           | /note="Organ: prostate; Vector: pDNR-LIB (Clontech);              |   |                     |
|                           | Site_1: Sfil (ggcgctcgctggc); Site_2: Sfil (ggcattatggc           |   |                     |
|                           | ); 5' and 3' adaptors were used in cloning as follows: 5'         |   |                     |
|                           | adaptor sequence: 5'-CACGGCCATTATGCGC-3' and 3' adaptor           |   |                     |
|                           | sequence: 5'-ATTCTAGAGCGCGCGCGGAGATG-dT(30)BN-3'                  |   |                     |
|                           | (where B = A, C, or G and N = A, C, G, or T). Average             |   |                     |
|                           | insert size 1.4 kb (range 0.5-4.0 kb). 14/15 colonies             |   |                     |
|                           | contained inserts by PCR. This library was enriched for           |   |                     |
|                           | full-length clones and was constructed by Clontech                |   |                     |
|                           | Laboratories (Palo Alto, CA)."                                    |   |                     |
| BASE COUNT                | 265 a   | 151 c   | 247 g               |
| ORIGIN                    |   |   | 192 t               |
| Query Match               | 11.6%   | Score 568.4   | DB 168; Length 855; |
| Best Local Similarity     | 89.5%   | Pred. No. 1.3e-143;   |                     |
| Matches 680; Conservative | 0;  | Mismatches 71;  | Indels 9; Gaps 6;   |
| OY                        | 3354  | ggaagtgatgacaccctcactatgagctggaagaagacggggaagagcctagatt       | 3413                |
| Db                        | 1   | GGAATGTGTGACACCTCTACATTTGAGCTGCAGAGAAAGCCGGGAAAGCCTAGATT      | 60                  |
| OY                        | 3414  | aagtatctgtglaaagaagacgaactgagatcttgatctgtgtcagaacattgtcaaggag | 3473                |

| Db         | 61  | AACTA-TGTGGTAAAGAAAGACTACGTGAGTA-TTGTCACAGCATGTGCAAGAGC                 | 118                                      |
|------------|---|---|--|
| QY         | 3474  | aatctcagaatcccgatctgtaagaagctgtaatcccaaggagagaccagatatatctgtgtgaatgggga | 3533                                     |
| Db         | 119   | AATTCCAGATGCCGATGTGAAGAGCTGATATCACAGGAGACACAGATATTAAATGATGGGGA          | 178                                      |
| QY         | 3534  | agaagcttgcgtaaatgctccccaagaagcggtgtccgctcttgtaagtgcttccccaagc           | 3593                                     |
| Db         | 179   | AGACGTTGCTAATGCCACCACCAAGAAAGCGGTGCCCTTTCGTTAAAGTCTCCACAGGCAC           | 238                                      |
| QY         | 3594  | agtaaccttggaaglttgtaagaatcaaaagctgtgtccattccatccagaagaggccatc           | 3653                                     |
| Db         | 239   | AGTAACCTTGGGAAGTGGGAGAGATCACTAAAGCTGTCATTCATTCCAGAGAGAGCCATC            | 298                                      |
| QY         | 3654  | tcaaacccagcccgaggagtgtaaggaagcgtctctcttccacccttccaccctctgagtc           | 3713                                     |
| Db         | 299   | TCAAAAGCAGCCAGGTGAGTGAAGGACGCTGTGTCATCTTTCACCTTTCACCTCTTGATC            | 358                                      |
| QY         | 3714  | cagtaaccttgatctacgtctggaagtagctcaagaagaatgcatctggcatctgtaataca          | 3773                                     |
| Db         | 359   | CAGTCACTCTGAGTCACTGGAAGTAGTCAAGAAGAAATGCATGGGATCTGAAATACA               | 418                                      |
| QY         | 3774  | gggatgaagaagatcgtgaatgtaaaaaggccctcctacgtaccctcacttgtaacagcatgc         | 3833                                     |
| Db         | 419   | GGGATTTAAGAACAGTCGAATATGAAAAAGGCGCTTACGTACTGCGGAATCACACATCGC            | 478                                      |
| QY         | 3834  | tggagagatgagcaaccacactctgtgtacgtgtcctatatltaatgcgaatgatccccaac          | 3893                                     |
| Db         | 479   | TGGAGGAGTAGGACACCCACTTGGTGATGTGCTATATTATTGAAATGATGACAC--AA              | 536                                      |
| QY         | 3894  | tggagttgcagcacagaccacaaaactcaaggtctgggagatgatgtatctacacatctgtg          | 3953                                     |
| Db         | 537   | CTGACTTTCAG--AAGACCCAAAATAATGAGTGGGGATAGGATTTGCACCATTTTGTGG             | 594                                      |
| QY         | 3954  | caactcaactctgggcatgactgactcacacccaagcaagtaactactgtaaaaatgcatctg         | 4013                                     |
| Db         | 595   | GACATTCCACAGTGAAGGATGATCAACAACCACAGCAAGTTAACTACTGAAAAATG--GTTTG         | 652                                      |
| QY         | 4014  | ctccatctgaatctcaggtgtgtctctgtgagagagctgtagtgtgtcacaaggtcaca             | 4073                                     |
| Db         | 653   | GTCCATTGAAATGACAGGTGGTGTGCGAGGAACATCGA-TGTTGTCCAGGTGCATCAGGA            | 711                                      |
| QY         | 4074  | ggagcctcgaatctcagtccttctcttcaactcgggctgacg                              | 4113                                     |
| Db         | 712   | GAGACCTGCAGATTCCAGTTTCTTAGTGGGGGAGGAG                                   | 751                                      |
| RESULT     | 11  |   |  |
| BE263698   | 701 bp  | mRNA  | EST                                      |
| LOCUS      | 6011921381  | NIH_MGC_7   | Homo sapiens cDNA clone IMAGE:3536024 5' |
| DEFINITION | mRNA sequence.  |   |  |
| ACCESSION  | BE263698  |   |  |
| KEYWORDS   | BE263698.1  | GI:9137243  |  |
| SOURCE     | human.  |   |  |
| ORGANISM   | Homo sapiens  |   |  |
| REFERENCE  | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; |   |  |
| AUTHORS    | Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.       |   |  |
| TITLE      | 1 (bases 1 to 701)  |   |  |
| JOURNAL    | NIH-MGC http://mhc.nci.nih.gov/                                   |   |  |
| COMMENT    | National Institutes of Health, Mammalian Gene Collection (MGC)    |   |  |
|            | Unpublished (1999)  |   |  |
|            | Contact: Robert Strausberg, Ph.D.                                 |   |  |
|            | Email: cgapds-remail.nih.gov                                      |   |  |
|            | Plate: L1CM217 row: 1 column: 09                                  |   |  |
|            | High quality sequence stop: 620.                                  |   |  |
| FEATURES   | Location/Qualifiers   |   |  |
| source     | 1..701  |   |  |
|            | /organism="Homo sapiens"  |   |  |
|            | /db_xref="taxon:9606"   |   |  |
|            | /clone="IMAGE:3536024"  |   |  |

```

/clone.lib="NH.MGC.7"
/cell_line="small cell carcinoma"
/lab_host="DH10B (phage-resistant)"
/Note="Organ: Lung; Vector: pPB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dr priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCAGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
```

```

BASE COUNT      215 a      144 c      194 g      148 t
ORIGIN
Query Match      11.4%; Score 554.6; DB 165; Length 701;
Best Local Similarity 99.1%; Pred. No. 7.1e-140;
Matches 568; Conservative 0; Mismatches 4; Indels 1; Gaps 1;
```

```

OY 1665 aattctatgtgctgcgaacatttgaaagatcacaataagcttgagcaacaac 1724
DB 67 AATTCTTATGTCCTGCGCAACATTGGAGAGATTCAAAATPAGCTTGGACAAACATC 126
OY 1725 tggagagatgagcactgatatcttctcatcactgacagacattccagaatt 1784
DB 127 TGGAGAGATGAGCAGCTGGATATTTTCTTCATACACTGGCAGACATTCAGAAAT 186
OY 1785 accagagcgaagaagagggatgaaagaagcgaactcaacaacacagatagcaa 1844
DB 187 ACCAGAGCGAAGAGAGGAGCGTGAAGAAAGCAACTTCAAAACACAGCATATAGCAA 246
OY 1845 ttgagatagcccaagcgggttgaaactctgagagagaaccaacaactcttgatcag 1904
DB 247 TTGGAATCAGCCCGGCGGCTGGAGAGAACCAACCAATCCTTAGGCATTCAG 306
OY 1905 catcttgagtcagagagatgagagtcgctgaagcaatgagaaagtatagagagat 1964
DB 307 CATTTGGTGGACGAGGATGGGAGTGGCTTAAGCAATGAGAAATATAGGGGCAAT 366
OY 1965 ttcatcaacaatgcttctgaaatagtcagctgcaaaaatgaaacttgaaactg 2024
DB 367 TTTCATCAAAACATGTTCTGAGAGATAGTCACACTGGCAAAATGGAACCTTGAACCTGG 426
OY 2025 agatagatcgttagagggcaaccacatcagtcagatcagagcagagaaaggtccatctg 2084
DB 427 AATATGAATCGTAGAGGCCACCCAGTCAGATCAGAGCCAGAGAAAGGCTCCATTTGTG 486
OY 2085 cagctgccccccaccctcctcagccttcgcaaatgggtagtgatcaacaacagtc 2144
DB 487 CAGTGTGCCCCCACCCTCTCTTCAGCCTTTCGCCAAATGGGATGATCAGACACAGATC 546
OY 2145 atctgcaagcaaaatctcacaagatctgagacaagaagagatgagttggtacagctgga 2204
DB 547 ACTGCAAGCAAAATCTCTCAAGATGTGAGCAAAAGAGATGAG-TTGGTTACAGCTGGA 605
OY 2205 aaatatcagagagcggttatgaaaccttaacagg 2237
DB 606 AATATTCAGAGAGCGTTATGAACTTAACAGG 638
```

```

RESULT 12
LOCUS      AV646633      672 bp      mRNA      EST      07-SEP-2000
DEFINITION AV646633 GIC Homo sapiens cDNA clone GLCAP10.3', mRNA sequence.
ACCESSION  AV646633
VERSION     AV646633.1 GI:9867647
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE   1 (bases 1 to 672)
```

```

AUTHORS      Qian,B., Wu,T., Huang,O., Kang,B., Gao,X., Xu,Z., Xiao,H.,
              Xu,X., Li,N., Peng,Y., Liu,F., Qu,J., Song,H., Cheng,Z., Qu,J.,
              Zeng,L., Xu,S., Gu,M., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M., Lu
              ,G., Yang,Y., Gu,Y., Chen,Z. and Han,Z.
TITLE        Homo sapiens cDNA clone
JOURNAL      Unpublished (2000)
COMMENT      Contact: Zeguang Han
              Chinese National Human Genome Center at Shanghai
              351 Guo Shouling Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
              201203, P. R. China
              Tel: 86-21-50801919(ex.45)
              Fax: 86-21-50801922
              Email: hanzg@chgc.sh.cn
              This clone is available at CHGC in Shanghai.
```

```

FEATURES
Source      Location/Qualifiers
1..672
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="GLCAP10"
/clone.lib="GIC"
/tissue_type="Corresponding non cancerous liver tissue"
/dev_stage="Adult"
/lab_host="SOLR"
/Note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2:
XhoI"
```

```

BASE COUNT      204 a      152 c      173 g      141 t      2 others
ORIGIN
```

```

Query Match      11.4%; Score 554.4; DB 32; Length 672;
Best Local Similarity 87.7%; Pred. No. 7.9e-140;
Matches 663; Conservative 0; Mismatches 8; Indels 85; Gaps 2;
```

```

OY 2711 gtggagtcataaagagcttaacagagcagtgagtagcagcagagtgagagacta 2770
DB 1 GTGGAGTCATCAATAAGACCTTACAGAGCATGAGGCTACACCGAGATGACGACTCA 60
OY 2771 aagtcggaatagatgactgctgctgtagatgaaatgttggttaccctatgaa 2830
DB 61 AAGTCGAGATGAGTACTGGCTGTGATGATGAATTTGTGTGTTACCTATTGAA 120
OY 2831 agttatagctcttgaaagacagcaagaatgacagtaaaacttcaactcactgaga 2890
DB 121 AGTTTATTAGCCTTCTGAGACAGCAAAATGACAGTAAACTTACATCCATGCTGAGA 180
OY 2891 atccagatcccaagcgtgttccttcaagcagctgtagcagctgagagaaaaaagaa 2950
DB 181 ATCCAGATTCCAGAGCTGTCTTCAGCAGCTGTGTCACGAGTGGAGAAAAAGAAC 240
OY 2951 gctccagctcttgatggtcccaagctgctgctcccaagaccgagatccatcgaata 3010
DB 241 GCTCCAGCTCTGTGATGTCACAGCTGTGCTCCCAAMCCGAGTCCATCCGAATA 300
OY 3011 caagcagatcatcaaacaccagcaatttctgctctgactcgaactgcccattatcc 3070
DB 301 CAAGCAGATCATCAACACGACGAATTTTGTCTTGTATCTTGCACCTGCCATTATCC 360
OY 3071 ctggctgcaaaacaaccatcgaagatttccaaaggcgcaacagggctgagcctgagatcg 3130
DB 361 CTGGCTGCAAAACACCATCGAGATTTCAAAGGGCGAAGAGGCTGGGCTGAGCATCG 420
OY 3131 ttgggggttcagacacgctgctggtgtgcttataatcactgaagtattatgaagaagag 3190
DB 421 TTGGGGGTTCAACACAGCTGCT-----
OY 3191 cagcatgtaaagatgaaagaccttggtgctgagatcagatcttagaggtgaatgag 3250
DB 443 -----GGTGAATGCAATTG 456
OY 3251 acttgaggaaggccacacatgataagcaatcaatgtctctgagacagcagcagagag 3310
DB 457 ACTTGAGGAAGGCCACACATGATGAAGCAATATGTCTGTGAGACAGAGCCACAGAGAG 516
```

QY 3311 tgcgcctgacacctctacagagatgaggcccatcaaaagaggagagtgctgtacaccc 3370  
 |||||||  
 Db 517 TGCCGCTGACACTCTACAGAGATGAGCCCATCAAAAGAGAGAGAGTGTGTACACCC 576  
 |||||||  
 QY 3371 tcaattatgagctgcagagaagccgggaaaggcctagagattagattgtgt-aaa 3429  
 |||||||  
 Db 577 TCACCTATTGAGCTCAGAGAAAGCCGGGAAAGCCCTAAGATTAAGTATGTGTGTAATA 636  
 |||||||  
 QY 3430 agaaacgactgagatattgtctcagacattgtc 3465  
 |||||||  
 Db 637 AGAAACGATCTGAGATTTTGGGGTAGACAAATGTC 672  
 |||||||  
 RESULT 13  
 AV646631 670 bp mRNA EST 07-SEP-2000  
 LOCUS AV646631 GLC Homo sapiens cDNA clone GICAPB07 3', mRNA sequence.  
 DEFINITION AV646631  
 ACCESSION AV646631 GI:9867645  
 VERSION AV646631.1 GI:9867645  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 670)  
 AUTHORS Qian,B., Wu,T., Huang,Q., Huang,C., Kang,B., Gao,X., Xu,Z., Xiao,H.,  
 Xu,X., Li,N., Peng,Y., Liu,F., Qu,J., Song,H., Cheng,Z., Qu,J.,  
 Zeng,L., Xu,S., Gu,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M., Lu  
 G., Yang,Y., Gu,Y., Chen,Z. and Han,Z.  
 TITLE Homo sapiens cDNA clone  
 JOURNAL Unpublished (2000)  
 COMMENT Contact: Zeguang Han  
 Chinese National Human Genome Center at Shanghai  
 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai  
 201203, P. R. China  
 Tel: 86-21-50801919 (ex.45)  
 Fax: 86-21-50801922  
 Email: hanzg@chgc.sh.cn  
 This clone is available at CHGC in Shanghai.  
 FEATURES  
 SOURCE  
 1. 670  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="GICAPB07"  
 /clone\_1b="GLC"  
 /tissue\_type="corresponding non cancerous liver tissue"  
 /dev\_stage="Adult"  
 /lab\_host="SOLR"  
 /note="Vector: pBluescript SK(-); Site\_1: EcoRI; Site\_2:  
 XhoI"  
 BASE COUNT 202 a 153 c 170 g 144 t 1 others  
 ORIGIN  
 Query Match 11.2%; Score 546.4; DB 32; Length 670;  
 Best Local Similarity 87.2%; Pred. No. 1.2e-137;  
 Matches 658; Conservative 0; Mismatches 12; Indels 85; Gaps 2;  
 QY 2711 gtgagatcatataaagagcttaacagagcagtggtgagcagcagatgagactca 2770  
 |||||||  
 Db 1 GTGAGATCATATATAAGAGCTTAACAGAGCATGGGTRAGCAGCCAGATGAGACTCA 60  
 |||||||  
 QY 2771 aagtcgagatcagatctggtcgtagatgaatgtgtgtgtgtacccatgaaa 2830  
 |||||||  
 Db 61 AAGTCGAGATCAGATCACTAGCTGTAGATGAATGTGTGTGTACCCATTGTA 120  
 |||||||  
 QY 2831 agttattagcctctcgaagacagcagagatgacagtaaaacttacatccatgctaga 2890  
 |||||||  
 Db 121 AGTTATTAGCTTCTGAAGACAGCAAAAGATGACAGTAAACTTACATCATCTGAGA 180  
 |||||||  
 QY 2891 atcagagttccagagctgttccctcagcagctggtgacagcagtgagaaaaaagaca 2950  
 |||||||  
 Db 181 ATCCAGATCCAGAGCTGTCTCTACAGAGCTGTGAGCCAGTGTGAGAAAAAAGACA 240  
 |||||||

QY 2951 gtcccaagctctctatgtgtccacagctctggtctccccaagacccggagttccatccgaata 3010  
 |||||||  
 Db 241 GCTCCAGTCTCTATGTGTGCCACAGCTGTGGCTCCCAAGAACCGAGATCTATCCGAATA 300  
 |||||||  
 QY 3011 caagcagatcacaacacagcaattttgtctctgtatccctgcaacctgccccatcc 3070  
 |||||||  
 Db 301 CAACGATCATCAACACACCAATTTTGTCTGTGATCCCGCAACCGCCCATTTATCC 360  
 |||||||  
 QY 3071 ctgagctggaacacacatctgagatttccaaaggcgagacgggctggtctgagctgc 3130  
 |||||||  
 Db 361 CTGGCTGCGAAACCAACATGAGATTTCCAAAGGCGAAGCGGCTGGCTGTGAGATCG 420  
 |||||||  
 QY 3131 ttgggggttcagacagcgtctggtccttattatccatgaagattatgaagaagag 3190  
 |||||||  
 Db 421 TTGGGGGTTGACAGACGCTGCT----- 442  
 |||||||  
 QY 3191 cagcatgtaaagatgaaagactctggtctgagatcagatcttagagggtgaatgtg 3250  
 |||||||  
 Db 443 -----GGTGAATGGAATTC 456  
 |||||||  
 QY 3251 acttgaggaagccacacatgtatgaagcaatcatgtctctgagacagagccacagagag 3310  
 |||||||  
 Db 457 ACTTGAGAAAGCCACACATGATGAAGCAATCAATGCTCTGAGACAGACGCCACAGAGAG 516  
 |||||||  
 QY 3311 tgcgcctgacacctctacagagatgaggcccatcaaaagaggagagtgctgtacaccc 3370  
 |||||||  
 Db 517 TGCCGCTGACACTCTACAGAGATGAGCCCATCAAAAGAGAGAGTGTGTACACCC 576  
 |||||||  
 QY 3371 tcaattatgagctgcagagaagccgggaaaggcctagagattagattgtgt-aaa 3429  
 |||||||  
 Db 577 TCACCTATTGAGCTCAGAGAAAGCCGGGAAAGCCCTAAGATTAAGTATGTGTGTAATA 636  
 |||||||  
 QY 3431 gaaacgactgagatattgtctcagacattgtc 3465  
 |||||||  
 Db 636 GAAACGATCTGAGATTTTGGGGTAGACAAATGTC 670  
 |||||||  
 RESULT 14  
 AV646666 670 bp mRNA EST 07-SEP-2000  
 LOCUS AV646666 GLC Homo sapiens cDNA clone GICAPB08 3', mRNA sequence.  
 DEFINITION AV646666  
 ACCESSION AV646666  
 VERSION AV646666.1 GI:9867680  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 670)  
 AUTHORS Qian,B., Wu,T., Huang,Q., Huang,C., Kang,B., Gao,X., Xu,Z., Xiao,H.,  
 Xu,X., Li,N., Peng,Y., Liu,F., Qu,J., Song,H., Cheng,Z., Qu,J.,  
 Zeng,L., Xu,S., Gu,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M., Lu  
 G., Yang,Y., Gu,Y., Chen,Z. and Han,Z.  
 TITLE Homo sapiens cDNA clone  
 JOURNAL Unpublished (2000)  
 COMMENT Contact: Zeguang Han  
 Chinese National Human Genome Center at Shanghai  
 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai  
 201203, P. R. China  
 Tel: 86-21-50801919 (ex.45)  
 Fax: 86-21-50801922  
 Email: hanzg@chgc.sh.cn  
 This clone is available at CHGC in Shanghai.  
 FEATURES  
 SOURCE  
 1. 670  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="GICAPB08"  
 /clone\_1b="GLC"  
 /tissue\_type="corresponding non cancerous liver tissue"  
 /dev\_stage="Adult"  
 /lab\_host="SOLR"



```

Db 457 ACTTGAGAAAGGCCACACATGATGAAAGCAATCAATGTCCTGAGACAGAGCCACAGAGAG 516
OY 3311 tgcgacctgacactctacagagatgaggcccaatacaagaaggaagtggtgacacc 3370
    |||||
Db 517 TGGCCTGACACTCTACAGAGATGAGCCCATCCAAAGAGAGAGAGTGTGTGACACCC 576
    |||||
OY 3371 tcaactatgagctgcagaagaagccgggaagaagcctagaataagtaattgtgtaaaa 3430
    |||||
Db 577 TCACTATTGAGCTGCAGAAAGCCGNNAAA-GCCTANGATTAAAGTATGTGTGCT-NAA 634
    |||||
OY 3431 gaaacgatctggagtaattgtgtcagacattg 3463
    |||||
Db 635 GAAACGATACTGAGGT-TTGTGTGACACATGG 666
    |||||

```

Search completed: July 12, 2001, 19:58:27  
 Job time: 8450 sec

7